UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

A landscape genomic approach for investigating growth performance in South African Bonsmara cattle

by

Charné Visser

Submitted in partial fulfilment of the requirements for the degree.

Magister Scientiae Agriculturae

MSc Agric Animal Science: Animal Breeding and Genetics

Department of Animal Science
in the Faculty of Natural and Agricultural Sciences
University of Pretoria
Pretoria

May 2023

Supervisor: Prof. Este van Marle-Köster
Co-supervisor: Prof. Carina Visser

Declaration

I, Charné Visser, hereby declare that this thesis, submitted for the MSc(Agric) Animal Science: AnimalBreeding and Genetics degree at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at any other University.

Charné

Visser

Pretoria

May 2023

Acknowledgments

First and foremost, all glory and honour to the Good Lord for graciously sustaining me throughout this endeavour. His loving kindness and mercy endure forever.

This project was not easy. There seemed to be challenges around every corner which made it difficult to persevere to the end. However, I owe the endurance to those that helped, guided, silently cheered, prayed, shared words of encouragement, reignited my interest in the topic through the expressions of utter fascination on their faces when I explained the project to them, and sacrificed time to assist when I could not understand the world of coding. There are too many to list by name, but if I ever spoke to you about my project, know that you are included in that list, and I am ever grateful to you for simply lending your ear.

The project would not have been possible without a very important group of people. Firstly, my supervisor and co-supervisor; Prof. Este van Marle-Köster and Prof. Carina Visser. Thank you both for your guidance, patience, and support over the last three years. Although the hiccups were frustrating to all three of us at some points, you both believed in the viability of the project when I could not see the bigger picture. I learned many academic related things from you both over this time, but I think I learned equally as much from you about strength, leadership, determination, hard-work, and standing firm in one's belief in oneself and one's ideas.

Then to my parents, thank you for your endless love, support, and (to dad) sponsorship of my studies. You two were there for me during all the behind-the-scenes struggles. You listened, you encouraged, you gave advice, and you told me when I needed to just suck it up, all from a distance that I think broke us all at times. I will never be able to thank you enough. I love you more than words can say.

Dr. Simon Lashmar, thank you for making my life so much easier than I was making it for myself. Without your assistance and guidance with some of the coding and analysis, I probably would have still been busy trying to figure some of it out. Thank you for your selflessness and for always being willing to help, despite how busy you are. Thank you also for the constant reassurance that things would be ok when I felt as though there was no end in sight. I knew when you walked into the lab that it would be a good day, because there would either be a break-through with my coding issues, words of encouragement, or a pop-culture reference that would keep a smile on my face for the remainder of the day. You are a light.

Dr. Khanyisile Hadebe, you came in at the last minute and practically saved the day. Thank you for your willingness to help a student you have never met, amidst your busy schedule. You provided the Profs. and I with the context and perspective that brought all the puzzle pieces together. Thank you for your patience when you explained difficult concepts and methodologies to me and thank you for answering all of our questions.

Jason Reding, thank you for helping me out at the very beginning and introducing me to the world of Plink. I am ever grateful for your patience in teaching me how to execute a QC and for answering all my many questions about very basic concepts. Thank you as well for your assistance with the genotypic data that I received from SA Stud Book and RMRD.

Dr. Marja Mostert-O'Neill, thank you for your guidance and assistance with the project. You helped me understand the premise of Landscape Genomics in a way that the articles could not, and you highlighted important factors that my supervisors and I were not initially aware of, ultimately steering us in the right direction. Thank you for the time you dedicated to our meetings and for your effort in sending me information about additional resources. You were instrumental in the successful completion of this project, thank you!

Lastly, a very big thank you to all my friends who uplifted, encouraged, and supported me, especially during this last year.

"The steadfast of mind You will keep in perfect peace, because he trusts in You." - Isaiah 26:3

Coram Deo

Abstract

Landscape genomics is used to investigate the potential association between the genotype of an animal and a specific environment. It is a relatively new approach applied in livestock genomics. In this study ŞA Bonsmara cattle from three provinces and different climatic regions were studied using a landscape genomics approach. The overall aim was to investigate potential association between the genotypes, growth, and the environment. Genotype and growth trait data from 4679 Bonsmara cattle were obtained for analysis. The cattle were grouped according to province, ownership, population size per province, and sex. After editing and pruning, the final list of animals included 766 cows from the Eastern Cape (418), Free State (224), and North-West (124) provinces. The genotypic data originated from four SNP array panels; GGP 80k (GeneSeek Genomic Profiler™), GGP 150k (GeneSeek Genomic Profiler™), IDB version 3 (International Beef and Dairy), and VersaSNP 50k (Weatherbys Scientific). The population structure of the cows was analysed through PCA plots and admixture plots, using GCTA64 and ADMIXTURE software respectively. The common SNPs across these panels were identified and quality control was conducted with PLINK; 25272 SNPs remained for downstream analysis. Weather data for the three provinces included summer and winter month temperatures, relative humidity, and average annual precipitation, from 2016 to 2021. Landscape genomics analysis was conducted on the weather variables and the 25272 common SNPs, using the latent factor mixed model (LFMM) landscape ecology association (LEA) software package in RStudio. Nine of the genes identified from analyses were previously reported to be associated with growth performance and adaptation. *HIBCH*; *CDH18*; *ATG7*; *GTDC1; MAP4K3*; *ADRA1A*; *PRKG1*; and *CSMD3* were previously confirmed for association with various growth performance traits, while *KCNJ16* and *PRKG1* were previously found to be associated with adaptation traits. A genome wide association study (GWAS) was consequently conducted to identify candidate loci associated with eighteen-month weight (18MW). Four SNPs were identified from the GWAS. However, no common SNPs were observed when these results were compared to those from the LEA landscape genomics analysis. Further studies on larger and more informative data sets will be needed for confirmation of the LEA results.

Table of Contents

List of Abbreviations

18MW	18 Month weight
AI	Artificial insemination
ADG	Average daily gain
AFC	Age at first calving
AFLP	Amplified fragment length polymorphism
BCS	Body condition score
BGP	Bovine genome project
BLUP	Best Linear Unbiased Prediction
BW	Birth weight
BAYENV	Bayesian method
Compno	Weather station reference number
DMI	Dry matter intake
DNA	Deoxyribonucleic acid
EBV	Estimated breeding value
EC	Eastern Cape
ECF	East Coast Fever
ENSO	El-Niño Southern Oscillation
FS	Free State
FCE	Feed conversion efficiency
FST	Fixation index
FAnGR	Farm animal genetic resources
GP	Gauteng province
GS	Genomic selection
GDP	Gross domestic product
GDS	Graphic data system
GEA	Genomic environmental association
GIS	Geographic information system
GVP	Gross value of production
GWAS	Genome-wide association study
GxE	Genotype by environment interaction
HapFLK	Haplotype-based extension of FLK
HWE	Hardy-Weinburg equilibrium

ID	Identification
IBD	Identity by descent
ICP	Inter-calving period
IMF	Intramuscular fat
KZN	KwaZulu-Natal
L	Limpopo
LD	Linkage disequilibrium
LH	Luteinizing hormone
LEA	Landscape ecology association
LFMM	Latent factor mixed model
LTHA	Long-term heat adaptation
MP	Mpumalanga Province
MW	Mature weight
MAF	Minor allele frequency
MAS	Marker assisted selection
NC	Northern Cape
NW	North-West
PC	Principal component
PCA	Principal component analysis
QC	Quality control
QTL	Quantitative trait loci
RDA	Redundancy analysis
RFI	Residual feed intake
RAD-seq	Restriction-associated digest DNA sequencing
Rhum	Relative humidity
RMRD	Red meat research and development
ROH	Runs of homozygosity
SA	South Africa
SC	Scrotal circumference
SAM	Spatial analysis method
$SAM_{(2)}$	Southern Annual Mode
SIOD	South Indian Ocean subtropical dipole
SMT	Summer month temperature
SNP	Single nucleotide polymorphism

sNMF	Non-negative matrix factorization algorithms
STHA	Short-term heat acclimation
THI	Temperature-humidity index
Tn	Maximum temperature
Tx	Minimum temperature
WC	Western Cape
WW	Weaning weight
WGS	Whole-genome sequencing
WMT	Winter month temperature
YW	Yearling weight

List of tables

Chapter 1

Chapter 2

Chapter 3

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 1. Introduction and Literature Review

1.1 Introduction

The South African agricultural industry contributes 2.4% to the country's gross domestic production (GDP) (Bureau for Food and Agricultural Policy (BFAP), 2021), of which the beef industry, contributed 0.12% to the gross value of production (GVP) during 2018 to 2020 (BFAP, 2021). Some of the predominant beef cattle breeds included in this industry are the Bonsmara, Tuli, Boran, Brahman, Simmentaler, Hereford, Afrikaner, Drakensberger, and Nguni breeds (van Marle-Köster *et al.*, 2021). These breeds are diverse in classification, varying between exotic, composite, and indigenous (Abin *et al.*, 2016). The Bonsmara is classified as a medium framed, composite breed and it was developed from crossbreeding *Bos taurus* and Sanga cattle. The distributions of the breeds used for its development were 5/8 Afrikaner, 3/16 Hereford, and 3/16 Shorthorn (Bonsma & Bonsma, 1985). The Bonsmara breed was developed specifically for the challenging climatic conditions of South Africa. It has grown to be the most popular beef breed in the industry (Brand *et al.*, 2021). The three primary South African beef production systems include weaner (cow-calf), long yearling, and two-year old systems (Govender, 2019). The weaner system is the most popular as farmers are able to reduce extended feed and/or medical costs by selling the weaners to feedlots or abattoirs early (Chadyiwa & Wepener, 2021). Beef farming in South Africa is mainly practiced extensively, on natural grazing with feedlot finishing being common (Brand *et al.*, 2021). Therefore, these systems are climate-dependent and susceptible to ongoing climate change. Farming with cattle that can tolerate and adapt to the variable climatic conditions is critical to ensure consistent animal welfare, production efficiency, and subsequently, stable economic returns (Chadyiwa & Wepener, 2021).

Climate often varies between geographic regions within a country, resulting in variation of genetic adaptation to certain climates within animals of the same breed (Vallejo-Trujillo *et al.*, 2022). Livestock can better utilize the natural resources of the environment in which they are kept, if they are well-adapted to that environment (Foster *et al.*, 2009). The majority of South African production environments are classified as semi-arid (Webb *et al.*, 2017). However, each of the country's nine provinces differ from one another in climate and/or vegetation. The eastern coastal provinces, KwaZulu-Natal (KZN), and the Eastern Cape (EC), present with sub-tropical climates and year-round precipitation. The Western Cape (WC) differs from all the other provinces (except for a few western regions in the Northern Cape (NC)) by having winter rainfall. The NC is classified as a semi-desert biome and has the lowest average annual precipitation compared to the other nine provinces (Climate: Northern Cape, 2023). Most of the inland provinces, Gauteng (GP), Free State (FS), North-West (NW), and Mpumalanga (MP), are relatively similar in climate. The GP and NW provinces experience milder winters (average of 13.35°C) whereas, FS and MP, experience colder winters (average of 8°C). All four

provinces have relatively warm summers, averaging to 28.25ºC. The Limpopo province (L) has mild winters and very hot summers (exceeding 40ºC). MP experiences higher than average annual precipitation in comparison to the other four inland provinces (South African Weather Service, 2020; Climate in Gauteng South Africa, 2023; Climate: Limpopo, 2023; Climate in Mpumalanga South Africa, 2023; Climate in North-West South Africa, 2023; Climate in Orange Free State South Africa, 2023). 70% of the national beef cattle herd is accounted for by Free State, North-West, Eastern Cape and KwaZulu-Natal provinces (Pienaar *et al.*, 2019). Beef production can be directly or indirectly influenced by the environment through its effect on cattle physiology or effects on feed sources. The subsequent phenotypic expression of adaptability and production traits can indicate whether the environmental influence on the cattle is positive or negative. Additionally, the phenotypic expression will provide insight into whether selection for such traits will lead to genetic progress or hinder it.

In the livestock sector, animal recording and genetic evaluations contribute to genetic progress by means of selection. Traditionally, the information required for genetic evaluations have been based on two types of information: phenotypic performance and pedigree records (Brand-Williams & Hayes, 2020). Best linear unbiased prediction (BLUP) calculations followed, providing breeders with estimated breeding values (EBV) for application in selection (Meuwissen *et al.*, 2001). The introduction of a third form of information, polymorphisms in the genome, has provided additional tools to improve the accuracy of genetic evaluations (Brand-Williams & Hayes, 2020). Genomic information is now included in genetic evaluations for genomic breeding values with higher accuracy. Genomic information yields the potential of improved breeding value prediction of younger animals, which may not have extensive records, by using single nucleotide polymorphisms (SNP). SNPS are available in commercial arrays for genotyping large numbers of animals (Garrick, 2011). Further, genomics has also created opportunities for the exploration of environmental influence on the genome and subsequently, the impact it could have on genetic progress, through landscape genomics.

Landscape genomics is an analysis method that combines spatial analysis and genotypic information to identify candidate genes that could be responsible for adaptive ability to a specific environment or productive performance in a specific environment (Joost & Negrini, 2010). Landscape genomics, founded through the geography sector, has been developed and applied since the early 2000s (Li *et al.*, 2017). It stems from *landscape genetics*, which was developed to evaluate the genetic adaptation of plants to a particular environment (Joost & Negrini, 2010). Landscape genomics was also developed for this purpose, but it has the potential to be successfully applied in the livestock industry, specifically to investigate areas in the genetic realm of livestock adaptation that have been difficult to quantify. The possibility of application and success of landscape genomics studies in the livestock industry is plausible as the input data required for analysis – spatial data, genetic data, environmental data – exists for most livestock animals (Storfer *et al.*, 2018). Pressure is placed on the livestock industry to pursue improved efficiency and sustainable production. The environmental adaptations that can be detected

from landscape genomics could aid those pursuits (Pariset *et al.*, 2012; Li *et al.*, 2017; Mdladla *et al.*, 2017; Webb *et al.*, 2017; Flori *et al.*, 2019; Grummer *et al.*, 2019). The genomic tool presents an opportunity to explore beef cattle DNA for markers associated with local adaptability. The observed variation in allele frequencies is 3tilization3 to be a result of functions of climatic, geographic, and disease and pathogen profiles (Mdladla *et al.*, 2017). When landscape *genetics* was initially introduced, the genetic data utilised for analysis consisted of microsatellite markers. The transition to landscape *genomics* resulted in the utilization of SNP markers for analysis (Storfer *et al.*, 2018).

1.1.1 Motivation

Growth performance is of interest due to its association with functional efficiency and the premise of selecting for adapted animals to improve reproduction and production potential (Bonsma, 1983). Previous studies linking bioregion to production and reproduction efficiency (Visagie, 2012; Webb *et al.*, 2017) warrant further investigation into the degree of such association. Flori *et al.* (2019) reported that Genomic Environmental Association (GEA) analyses facilitated the detection of genetic variants associated with various environmental variables, indicating a degree of genetic adaptation to a specific environment. Breeding with animals that are well-adapted to their environment and regional climate is important because the consequences of poor adaptation can be observed through diminished immune response, metabolic efficiency, and reproductive performance (Nardone *et al.*, 2010; Krebhiel *et al.*, 2019).

Climate change predominantly influences livestock production and performance through changes in feed resources and ambient temperature (Cheng *et al.*, 2022). Feed resources indirectly affect livestock productivity, the distribution of livestock diseases and parasites, the sustainability and buffering ability of ecosystems, and rangeland carrying capacity (Rust & Rust, 2013). Further, abnormally high or low ambient temperatures stimulate physiological responses in livestock that tend to hinder growth performance (Furstenberg & Scholtz, 2009).

Landscape genomic analysis has been effective in identifying candidate genes associated with adaptation to environmental factors (Frichot & Francois, 2015a). Landscape genomic analysis can also play an integral role in predicting population survival by facilitating the analysis of intricate Genotype by Environment (GxE) interactions that influence population abundance, distribution, and growth rates (Grummer *et al.*, 2019). There is an ever-growing need for conservation of livestock genetic diversity, especially in the present time when adaptation to changing climatic conditions is vital (Pariset *et al.*, 2012). By combining spatial analysis and genomics, landscape genomics could provide valuable insight into the extent that a local environment and its geography can influence the genetic structure of livestock populations (Pariset *et al.*, 2012). Such information may allow for more purposeful selection of genetic resources that will survive and perform in a future greatly affected by global warming and striving for food security.

1.1.2 Aim of the study

The aim was to investigate growth performance of Bonsmara cattle in the Eastern Cape, Free State, and North-West provinces of South Africa, using a landscape genomic approach. The Free State, North-West, and Eastern Cape provinces are being exclusively investigated in the study due to the available numbers of registered Bonsmara cattle compared to the other six provinces.

Objectives:

To compile data sets for different geographic regions with climatic variables and phenotypes.

To compile dataset with genotypes to match the climatic-phenotypic data.

To process the genomic and environmental datasets to partition the genomic variation among geographic and climatic variables and to remove redundancies in environmental data.

To generate a list of candidate loci that contribute to the genomic adaptation of Bonsmara cattle to their respective environments.

To annotate the candidate loci to determine whether corresponding genes are associated with growth performance or adaptation.

1.2 Literature review

1.2.1 Introduction

Landscape genomics is the field of study whereby genotypic information from spatially referenced populations, environmental data from varying landscapes, and the relevant geographic data are combined to identify genomic regions involved in local adaptation (Mdladla *et al.*, 2018). Landscape genomics has value in the livestock sector due to the existence of animal ecotypes and biological types in different regions of countries.

In this review, a brief overview is provided on the factors influencing livestock production, traits of economic importance, and the role of Bonsmara cattle with reference to their selection for adaptation and production in various geographical regions. This is followed by a short discussion on genome-wide association studies (GWAS) and the application of SNPs with specific reference to landscape genomics and how it can be applied to the livestock industry.

1.2.2 South African environment and its effects on beef cattle performance

The genetic selection of cattle to meet specific breeding objectives becomes null and void if the selected individuals are not productive. Functional efficiency is a term originally coined by the late Professor JC Bonsma (Bonsma & Bonsma, 1985) and relates to reproduction, adaptability, and frame. The importance of functional efficiency is that selection of the most functional breeding animal based on environmental constraints, will prevent a detrimental decline in performance of the next generation (van der Westhuizen, 2019).

The late Professor Jan Bonsma outlined 16 environmental factors that influence the performance of cattle (Bonsma & Bonsma, 1985). These factors were visually arranged in the format of a wheel and entitled the Livestock Ecology Wheel (Figure 1.1). The environmental factors of importance that relate to this study are temperature, rainfall and humidity, and nutrition.

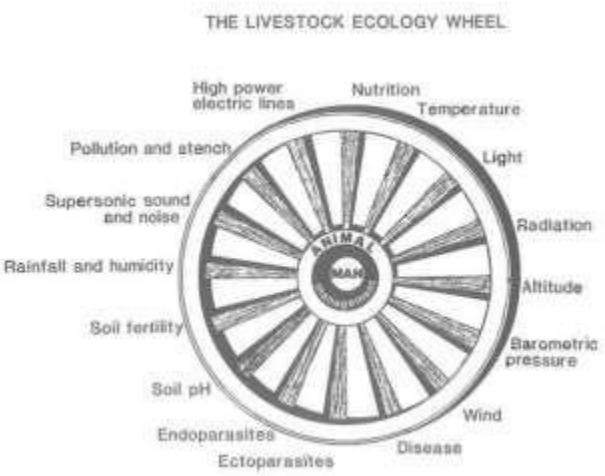


Figure 1.1 The Livestock Ecology Wheel developed by Prof. Jan Bonsma (Bonsma & Bonsma, 1985).

Natural vegetation in South African provinces

South Africa is known for its vast variety of flora and fauna, as well as the distinctiveness of vegetation species and quality found in each province. Most South African bioregions are arid or semi-arid, with high ambient temperatures that may adversely affect the production of livestock during summer (Du Preez *et al.*, 1992; De Jager, 1993). Moreover, temperature, in tandem with rainfall, influences the vegetation availability, quality and distribution (Visagie, 2012). Natural veld is often the only feed source available to grazing cattle. Resultantly, poor nutrient intake, due to poor forage quality, is an important constraint in beef cattle production in South Africa (Visagie, 2012). This is a factor that ultimately influences the growth performance of cattle. Figure 1.2 illustrates which bioregions are found in the different provinces of the country and Table 1.1 summarises the relevant bioregion information for the Eastern Cape, Free State, and North-West provinces.

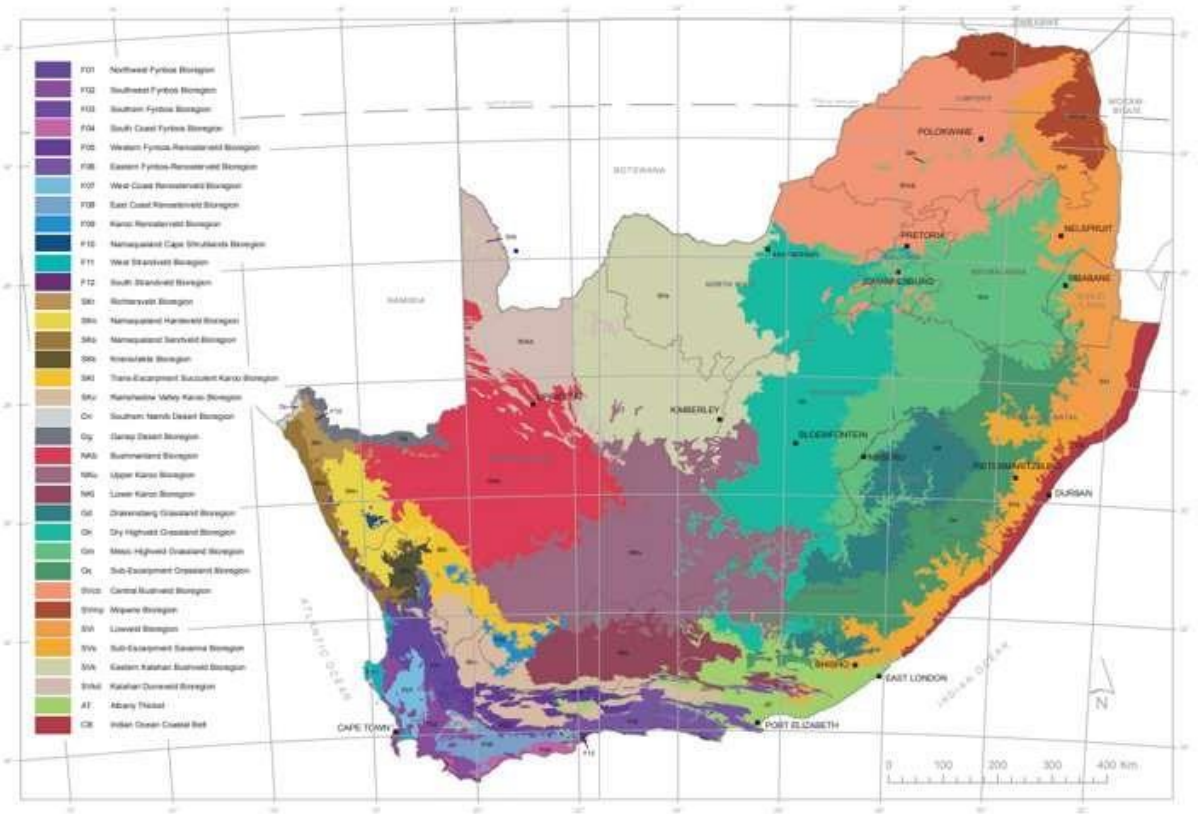


Figure 1.2 A map depicting the bioregions of South Africa (Mucina *et al.*, 2006).

The three primary categories of veld type are sweet, sour, and mixed veld, with subcategories comprising of combinations of the primary categories (Figure 1.3) (Tainton, 1999). Sourveld is found in areas with high rainfall (at least 650mm of precipitation) and soils that are more acidic. The vegetation in these areas becomes less palatable and nutritious once it reaches maturity (Ellery *et al.*, 1995; Van Rooyen, 2002). Sweetveld occurs in areas with low rainfall (200 to 500mm) and soils that are 6haracterized by a high base and mineral status. The vegetation is nutritious throughout the year granted the veld is managed properly (Ellery *et al.*, 1995; Van Rooyen, 2002). Further, sourveld is predominant

in areas where the carbon assimilation is high relative to the nutrient supply, and sweetveld occurs in areas where the carbon assimilation is low relative to the nutrient supply. Mixed veld is the intermediate of sour and sweetveld (Ellery *et al.*, 1995; Tainton, 1999). Figure 1.3 indicates that sourveld is more common in the eastern part of the country (where rainfall tends to be higher) and sweetveld is more common in the west (where rainfall tends to be lower).

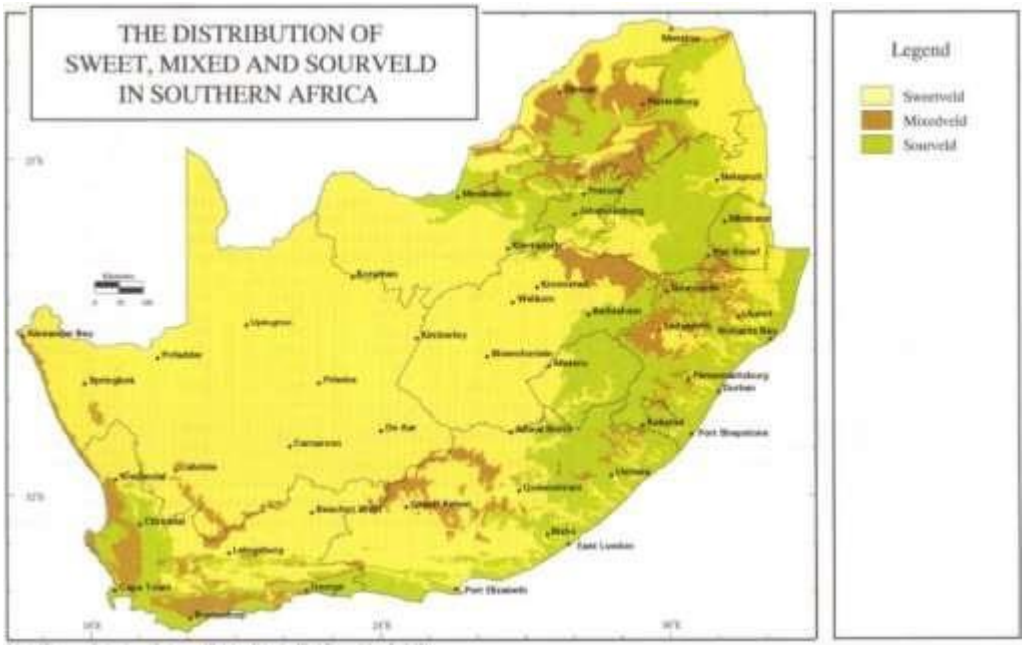


Figure 1.3 The distribution of sweet, mixed, and sourveld in South Africa (Tainton, 1999).

South Africa has a diverse climate ranging from temperate conditions in the Western and Eastern Cape to sub-tropical conditions in the Free State and North-West. The temperatures and precipitation levels vary throughout the country, resulting in variations of veld types as they are distributed through the grassland biome.

The major climatological factors that influence livestock production are precipitation and temperature (Hafez, 1968). The effects of changes to the climate in South African provinces are predominantly observed in alterations to feed resources, which subsequently influence the production and performance of livestock. Additionally, feed resources impact the distribution of livestock diseases and parasites, rangeland carrying capacity, the sustainability and buffering ability of ecosystems (Rust & Rust, 2013). One of the ways through which climate change can bring about modifications to these factors is when the primary productivity of crops, forages, and rangelands change, which consequently change the quantity of forage and fodder available for dry season feeding (Thornton *et al.*, 2007). Another feed resource alteration that climate change brings about is the change in species composition of rangelands, which may further implicate the types of animal species able to graze the land (Thornton *et al.*, 2007). Additionally, the digestibility and rates of degradation of plant species can be altered by climate change, as the plants themselves grow to adapt to a different climate (Thornton *et al.*, 2007).

Table 1.1 Summary of veld type and bioregion classification of Eastern Cape, Free State, and North-West provinces (compiled from Tainton, 1999 and Mucina *et al.*, 2006).

Eastern Cape	**Free State**	**North-West**
Well-rounded blend of sweet, sour, and mixed veld **Eight bioregions:** • Dry Highveld Grassland Bioregion, • Albany Thicket, • Eastern Fynbos-Renosterveld Bioregion, • Lower Karoo Bioregion, • Sub-Escarpment Savanna Bioregion, • Sub-Escarpment Grassland Bioregion, • Drakensberg Grassland Bioregion, • Upper Karoo Bioregion	Predominantly sweetveld with some mixed veld and sourveld in the north-eastern regions **Six bioregions:** • Mesic Highveld Grassland Bioregion, • Central Bushveld Bioregion, • Dry Highveld Grassland Bioregion, • Upper Karoo Bioregion, • Drakensberg Grassland Bioregion, • Eastern Kalahari Bushveld Bioregion	Predominantly sweetveld, some sour and sourish-mixed bushveld **Four bioregions:** • Dry Highveld Grassland Bioregion, • Eastern Kalahari Bushveld Bioregion, • Central Bushveld Bioregion, • Mesic Highveld Grassland Bioregion

Predicted changes to temperature (overall increase), humidity (overall increase), precipitation (overall decrease), and precipitation variability (overall increase) would be the driving forces behind negative impacts on livestock production (Angel *et al.*, 2018; Chadyiwa & Wepener, 2021). A 2017 study observed that from 1985 to 2014 in Gauteng and Limpopo, the average annual precipitation had decreased by 57.361mm and 17.014mm, respectively, and the average maximum temperature increased by 0.49 °C and 0.54 °C, respectively (Elum *et al.*, 2017). Such climate change can have direct implications on livestock growth, health, and reproduction.

When livestock struggle to dissipate excess heat, due to the ambient temperature, relative humidity, and radiant energy exceeding the normal parameters in which livestock can comfortably exist, they succumb to heat stress which can be detrimental to their production abilities (Daramola *et al.*, 2012; Rashamol *et al.*, 2020; Cheng *et al.*, 2022). Animals primarily transfer heat by evaporation as it is independent of a temperature gradient. However, since humidity affects the evaporation rate, the temperature humidity index (THI) becomes more relevant when high temperatures and high humidity persist (Rust & Rust, 2013).

Climate change will have the greatest impact on farming systems that are dependent on the environment, such as extensive grazing and mixed farming systems. Reduced precipitation corresponds to increased droughts, which can negatively impact poorly adapted crop varieties and result in shorter pasture growing periods (Konapala *et al.*, 2020; Cheng *et al.*, 2022). Additionally, a demand for more efficient water usage will result with reduced average annual precipitation, to meet the water requirements of

livestock and vegetation (Webber *et al.*, 2018; Cheng *et al.*, 2022). However, a risk associated with water, especially when it is obtained from sources other than natural precipitation, is that of chemical contaminants. Ground water located close to rural settlements or larger cities are at risk of containing organic or inorganic, biological, or heavy metal contaminants (Nardone *et al.*, 2010). Animals need two to three times more water when exposed to temperatures exceeding their thermo-neutral zones. Therefore, the risk of excessively consuming dangerous contaminants or water with altered pH, increases if such water is exclusively available to them. Effects may be observed in the metabolism, fertility, and digestion of the animals (Nardone *et al.*, 2010).

Temperature and relative humidity

Most South African beef cattle farms operate as extensive production systems where cattle are constantly exposed to natural conditions. Managing the cattle to reduce the climatic impact remains a challenge, especially when rapid changes to climatic conditions occur (Gaughan & Cawdell-Smith, 2015; Lees *et al.*, 2019). Studies investigating the influence of high environmental temperature and solar radiation on cattle production, reported reduced dry matter intake, which subsequently led to reduced average daily gain, and lower clean dressing percentages (Nardone *et al.*, 2010; Lees *et al.*, 2019). Solar radiation has an inhibiting effect on the thermoregulatory abilities of animals, especially in areas where natural shade is not abundantly available or artificial shading solutions are not possible (Lees *et al.*, 2019).

The effects of rising ambient temperatures on the performance of beef cattle include reduced feed conversion efficiencies, reproduction rates, and weight gain, and increased prevalence of internal parasite infections, and prevalence of vector-borne diseases (Summer *et al.*, 2019; Cheng *et al.*, 2022). Animals have a thermo-neutral zone, within which temperature is regulated through the control of sensible heat loss – there is no regulatory alteration to metabolic heat production or evaporative heat loss (Gonzalez-Rivas *et al.*, 2020; Cheng *et al.*, 2022). However, when environmental variables collectively push the thermo-neutral zones to their upper limits or beyond, the affected animal will begin to suffer from “heat stress” (Visagie, 2012; Cheng *et al.*, 2022). The animal’s body will physiologically respond to heat stress, to maintain the body temperature within normal range, by increasing heat loss and reducing heat production (Bernabuccil *et al.*, 2010; Rashamol *et al.*, 2020). This process of heat acclimation has two phases: short-term heat acclimation (STHA) and long-term heat adaptation (LTHA) (Collier & Zimbelman, 2007; Visagie, 2012). STHA begins during heat stress periods, and involves the alteration of cellular signaling pathways, which ultimately result in the reprogramming of cells to withstand the detrimental impact of heat stress (Horowitz, 2001). LTHA occurs after STHA and is characterized by the ultimate enhancement of metabolic processes and signaling pathway efficiencies. These improvements are triggered by the modified gene expression brought on by heat stress and cellular responses (Horowitz, 2001).

The temperature-humidity index (THI) is used as a thermal stress indicator for livestock. The threshold limits are as follows: values <74 are defined as 'alert'; 75<THI<79 are defined as 'danger'; and values exceeding 79 are classified as 'emergency' (Amundson *et al.*, 2006). In periods of heat stress, respiratory rates increase, and feed intake decreases (Foster *et al.*, 2009). Thus, in regions where the THI is high, the overall mature weights of animals may be lower than what is deemed "average" by breed standards, due to the consumption of less feed over their lifetimes.

Decreased feed intake is the primary response to heat stress observed in ruminants. When ambient temperatures are higher than normal, the normal metabolic heat released during the fermentation process of feed within ruminants consequentially compromises their thermoregulation ability (Gonzalez-Rivas *et al.*, 2020). However, this places the animals into a negative energy balance (Bernabuccil *et al.*, 2010; Visagie, 2012), which limits weight gain and negatively influences their mature weights. Additionally, in the extreme event of chronic heat stress, metabolic adaption comprising of alterations to endocrine function, basal metabolism, water and electrolyte metabolism, and rumen fermentation, may also take place (Padodara & Jacob, 2013; Angel *et al.*, 2018). Heat-stressed animals have much greater water requirements than thermoneutral animals. The main driver behind increased water intake is for thermoregulation to counter water evaporation that occurs from panting and sweating (Gonzalez-Rivas *et al.*, 2020) and to cool down the reticulum-rumen to lower the internal body temperature (Bewley *et al.*, 2008; Gonzalez-Rivas *et al.*, 2020).

Heat stress has been reported to compromise reproductive efficiency of livestock in tropical or subtropical regions, leading to a negative impact on the results of animal selection and eventual meat and milk production (Alves *et al.*, 2014; Angel *et al.*, 2018; Rahman *et al.*, 2018). Raised internal body temperatures compromise oocyte growth in cows by altering the dynamics of hormone secretion during the oestrus cycle (Angel *et al.*, 2018). Most studies have observed that cows suffering from heat stress have diminished luteinizing hormone (LH) secretion and that which is secreted does not function properly (Bridges *et al.*, 2005; Wolfenson & Roth, 2019). Additionally, the intrauterine environment is compromised when high environmental temperatures persist because blood flow towards the uterus is decreased to direct increased blood flow to the extremities, to allow greater heat dissipation. Resultantly, the temperature in the uterus increases (Dash *et al.*, 2016). Heat stress has also been associated with increased embryo mortality and defective embryo development in cattle (Angel *et al.*, 2018).

Another influence of high environmental temperatures on reproduction is reduced pregnancy rate. A study reported reduced pregnancy rates when average daily environmental temperatures and average daily THI exceeded 16.7°C and 72.9, respectively (Amundson *et al.*, 2006). Lowered conception rates have also been reported to be associated with raised rectal and uterine temperatures (Gwazdauskas, 1985).

In bulls, higher environmental temperatures result in reduced semen concentration, number of spermatozoa, motile cells, and major sperm defects per ejaculate (Mathevon *et al.*, 1998; Rahman *et al.*, 2018). Although the impact of heat stress on sperm quality or ejaculate volume may not be immediate or extreme, prolonged exposure can affect field fertility on a long-term basis (Gwazdauskas, 1985). An in vivo study by Rahman *et al.* (2018) investigated the effects of prolonged heat stress on sperm quality and alterations in sperm chromatin conformation were observed. This led to increased sperm abnormalities and low fertility (Rahman *et al.*, 2018). Bulls exposed to high ambient temperatures (>29°C), were reported to present with decreased initial sperm motility, concentration and total sperm counts (Tucker & Oxender, 1980; Gwazdauskas, 1985). Further exposure of the bulls to extreme temperature (40°C), resulted in decreased motility and percent live spermatozoa, and increased percent of abnormal spermatozoa. (Tucker & Oxender, 1980; Gwazdauskas, 1985).

Some mitigation strategies and adaptation factors to combat the effects of heat stress and ultimately improve efficiency in the beef sector (greater output of reproduction, survival, and growth rates), include greater emphasis on efficient nutrition, continuous genetic improvement, and modifying production systems to diminish their carbon footprint (Gaughan *et al.*, 2019). Mitigating heat stress will not only improve animal well-being, but it will also aid in the prevention of economic losses due to poor production, fertility, and animal welfare (Gonzalez-Rivas *et al.*, 2020).

Precipitation

South African bioregions receive varying average annual precipitation. The inland provinces, GP, FS, L, NW, and MP receive between 400 to 620mm per annum and the eastern coastal provinces, EC and KZN, receive between 550 to 850mm per annum (Climate in Eastern Cape South Africa, 2023; Climate in Gauteng South Africa, 2023; Climate in KwaZulu-Natal South Africa, 2023; Climate in Limpopo South Africa, 2023; Climate in Mpumalanga South Africa, 2023; Climate in North-West South Africa, 2023; Climate in Orange Free State South Africa, 2023). KZN and parts of EC receive year-round precipitation, and MP has roughly two to three dry months in the year. GP, FS, L, and NW, however, do not receive winter precipitation, thus are at risk of long drought periods. Growth performance has been known to have a curvilinear relationship to the seasonal distribution of rainfall – which in turn determines the available vegetation of a specific season (Foster *et al.*, 2009). South Africa is a country subjected to precipitation variability, frequent droughts, floods, and high ambient temperatures (IPCC 2007, 2013). Water shortages and heat stress were found to be the two most impactful factors on cattle production in semi-arid environments (Dzavo *et al.*, 2019). Studies investigating the impact of water scarcity on livestock farming in Sub-Saharan Africa, reported that droughts cause cattle deaths, feed shortages, and weak immunity in cattle (Megersa *et al.*, 2014; Tolemariam *et al.*, 2015; FAO, 2016; Magita & Sangeda, 2017; Kimaro *et al.*, 2018).

Without sufficient precipitation, crops and natural veldt fail to grow and the resulting feed shortages negatively affect the cattle production and reproduction (Rojas-Downing *et al.*, 2017; Dzavo *et al.*, 2019). The vegetation that is available in low precipitation seasons is more likely to be of poor quality (Rojas-Downing *et al.*, 2017). Contrastingly, bioregions where high rainfall is more common may offer greater quantities of grazing, but it may be of lower nutrient quality due to likely being classified as sourveld (Webb *et al.*, 2017). Further, water intake aids feed intake; easing the digestion of dry feed (Rojas-Downing *et al.*, 2017).

Nutrition

Extensive farming systems in South Africa are associated with semi-arid environments. Since these systems are entirely dependent on natural resources, forage and concentrate resources can become limited in quantity and quality when environmental stress is severe and prolonged. Extended drought periods will negatively affect vegetative growth and forage water content, causing them to be unpalatable and highly fibrous (Konapala *et al.*, 2020). Fibrous and unpalatable forages are likely to reduce voluntary feed intake. High fibre feed may also result in excess heat production due to excessive fermentation, which could lead to an increased thermoregulatory demand for water (Polley *et al.*, 2013).

Forage quality and quantity has seasonal variability. The quality/digestibility determines the dry matter intake (DMI) and subsequently the nutrient intake. Low digestibility equates to increased fermentation time, which is not desirable as it results in increased metabolic energy expenditure with low energy return (Maas, 1987). Winter is generally associated with poorly digestible feed and summer with highly digestible feed (Maas, 1987). The natural veld in semi-arid, South African environments is comprised of rangeland that is poor in both quality and quantity. From a nutrition perspective this translates to malnutrition and consequently, poor immune status of animals (Dzavo *et al.*, 2019). Further, the rangelands in South Africa are classified as sour, sweet, or mixed according to the rainfall they receive and the vegetation they contain (Mapiye *et al.*, 2009). Rangelands receiving high annual rainfall (600 to 800mm), typically on the eastern side of the country, consist primarily of annual grass species and are classified as sour veld. These grasses lose nutritive value and palatability in the dry seasons (Ellery et al., 1995). Contrastingly, rangelands receiving less than 500mm of annual rain, more common on the western side of the country, consist primarily of perennial grass species and are classified as sweet veld (Ellery *et al.*, 1995). These grasses remain nutritious and palatable all year round but may be lower in quantity due to the lower rainfall.

Apart from the clear role that available feed plays in growth and production ability of cattle, it is also indirectly integral to efficient and successful reproduction. This is primarily due to body condition and mass being critical to the various stages of reproduction. Ideally, heifers should reach puberty at 14 to 16 months of age and first calve at around two years (Michael *et al.*, 2019). Environment and nutrition are two of the main factors that influence the age at which heifers reach puberty. Regarding nutrition,

energy intake and subsequent body weight are the influencing factors (Buskirk *et al.*, 1995). The goal is for heifers to have reached approximately 66% of their mature body weight by their first breeding season. Low energy intake will delay the onset of puberty in both heifers and bulls (Maas, 1987).

Further, it is essential that pre-partum energy intake is sufficient to allow for normal fetal growth and development, calf survivability and growth, postpartum breeding efficiency, and short inter-calving period (ICP) (Maas, 1987). Adequate energy intake is especially important during the third trimester of pregnancy as approximately 75% of fetal growth occurs during this time (Schoonmaker & Eastridge, 2013). When cows receive inadequate energy during this stage of gestation, reduced neonatal survivability and low calf birthweights are observed. Additionally, cows tend to produce poor quality colostrum when their feed intake does not meet their energy requirements during the final stages of gestation (Maas, 1987; Buskirk *et al.*, 1995). This can be detrimental to calf survival. Further, milk production also tends to be lower in malnourished cows, which ultimately leads to low calf growth rates and low weaning weights. (Michael *et al.*, 2019). The low body condition scores (BCS) of such cows will result in them having a slow return to estrus, reduced future conception rates, and longer inter-calving intervals (Maas, 1987; Michael *et al.*, 2019).

1.2.3 Traits of economic importance

The growth traits that are recorded for South African beef cattle are generally easy to measure on farm. Some of these economically important traits include birth weight (BW), weaning weight (WW), mature weight (MW), and average daily gain (ADG). In addition, fertility traits such as age at first calving (AFC), inter-calving period (ICP), and scrotal circumference (SC) (Abin *et al.*, 2016; Visser *et al.*, 2020). A summary of the heritability values for these traits are shown in Table 1.2. SA Stud Book utilises an animal recording system called Logix to store the performance records of the breeds which they render a service for. The performance records and pedigree records are required for genetic evaluations that utilise best linear unbiased prediction (BLUP) animal models to improve selection efficiency and obtain genetic change. Genetic evaluations allow for the calculation of estimated breeding values (EBVs) which enable accurate selection decisions about economically important traits and are useful in determining genetic trends (Abin *et al.*, 2016).

There are many factors that can alter the weight of an animal from birth to maturity, but growth curves, developed through an equation by Brody & Lardy (1946), as well as the generally, consistently strong genetic correlations (Table 1.2) between weights at different ages indicate that an individual's birth weight sets the foundation for its future growth performance (Thonney, 2015; Gathura *et al.*, 2020). Mature weight (h^2 = 0.56) is, genetically, predominantly influenced by additive genetic variation (Zimmermann *et al.*, 2021). A cow's growth performance plays an integral role in her maintenance requirement and the manner with which she responds to the environment (Kattnig *et al.*, 1993; James,

2009; Webb *et al.*, 2017). For these reasons, it can be deduced that cow size should also influence a cow's adaptive ability.

Almost 30 years ago, suggestions were made to improve beef cow efficiency by selecting for cow size based on the environment (Kattnig *et al.*, 1993). The optimal mature size is dependent on the production system and environment (Visagie, 2012). Research has suggested that selection for an optimal cow size should be directed towards the animals with mature sizes superiorly adapted to the breeding system, environment, and market factors of the regions they were produced in (Dickerson, 1970, 1978). Authors have made varying recommendations on which frame sizes would be optimal for which environments. Large-framed cattle were suggested to perform better in semi-arid tropics by some (Bonsma, 1983), and by others, in regions where forage is in abundant supply (Dickerson, 1978; Solis *et al.*, 1998). Additionally, smaller-framed cattle were thought to be at an advantage in humid tropics by some (Bonsma, 1983), and by others, regions with hot and dry climates (Dickerson, 1978; Solis *et al.*, 1998). Small to medium framed cattle breeds have been observed to perform more successfully in the hot and dry conditions of South Africa (Visagie, 2012). Small-framed breeds weight between 320 to 410kg at mature size (i.e., Nguni breed) and medium framed breeds weigh between 500 to 580kg at mature size (i.e., Bonsmara breed), with bulls being on the heavier side (Strydom, 2008; Brand *et al.*, 2021). By selecting for an optimal cow size, potential is created to improve the adaptive ability of beef cows. This should in turn improve the overall efficiency of beef production in diverse environments.

Table 1.2 Summary of heritabilities and genetic correlations for traits of interest for beef cattle.

Trait	**Heritability**	**Reference***
Direct birth weight	0.23-0.39	b, c, f, g
Weaning weight	0.07-0.27	b, c, f, g
Yearling weight	0.14-0.27	c, f, g
18 Month weight	0.29	f
Mature weight	0.32-0.56	b, e, f, g, h
Age at first calving	0.18-0.40	f, g
Inter-calving period	0.01-0.03	f, g
Scrotal circumference	0.40-0.44	b, g
Rectal temperature	0.18-0.22	d, g
Genetic Correlations		
Birth weight and weaning weight	0.44-0.53	a, g
Birth weight and mature weight	0.63	g
Weaning weight and yearling weight	0.84	g
Yearling weight and mature weight	0.43	
Rectal temperature and periodic weight gain	-0.20 to -0.49	a
Rectal temperatures and birth weight, mature	-0.05 to -0.20	a

weight, and dry season gain	

*Burrow, 2001[a]; Maiwashe *et al.*, 2002[b]; Van Graan *et al.*, 2004[c]; Prayaga *et al.*, 2009[d]; Crook *et al.*, 2010[e]; Van der Westhuizen *et al.*, 2011[f]; Gathura *et al.*, 2020[g]; Zimmermann *et al.*, 2021[h]

1.2.4 Bonsmara breed

The Bonsmara beef cattle breed was developed at the Mara Research station in South Africa between 1937 and 1963 to produce a breed that would be well adapted to and perform productively in the diverse climate regions of South Africa (Bonsma, 1980; van Marle-Köster *et al.*, 2021). The breed is comprised of 5/8 Afrikaner, 3/16 Hereford, and 3/16 Shorthorn breeds, and is thus a blend of European taurine, African taurine, and African indicine (Sanga) breeds (Mostert & Exley, 2000; van Marle-Köster *et al.*, 2021). Mature cows weigh approximately 500 to 550kg (Brand *et al.*, 2021).

The SA Bonsmaras have been selected to produce economically in sub-tropical climates (Bonsma & Bonsma, 1985). Functional efficiency played a major role in breed development. To be registered as stud animals, Bonsmara cattle are required to be screened for functional efficiency by breed inspectors (Webb *et al.*, 2017), and animals found with structural defects or to be functionally inefficient are culled. Some of the characteristics that have aided in the breed's success in the South African beef industry are their adaptability on veld, their growth under both intensive and extensive conditions and their mothering ability; which were the traits targeted in the Afrikaner, British Hereford, and Milk Shorthorn during the breed's development (Bonsma, 1980; van Marle-Köster *et al.*, 2021). The breed's popularity has allowed it to grow into the most prevalent beef breed in South Africa (Webb *et al.*, 2017; Brand *et al.*, 2021).

The traits of interest to this review recorded with Logix for Bonsmaras are birth weight (BW), weaning weight (WW), yearling weight (YW), and 18-month weight (18MW). The phenotypic averages of these traits are 35.21kg, 228.05kg, 271.16kg, and 356.27kg respectively (SA Stud Book/Logix data, 2021). It has been observed that Bonsmara breeders typically assume that certain types or sizes of cattle are better adapted to specific production regions in South Africa (Webb *et al.*, 2017). Phenotypic performance for mature cow weight will be influenced by genetic and environmental effects, and GXE associations can also be influential (Neser *et al.*, 2008; Garrick & Enns, 2010; Jordaan, *et al.*, 2021). Bonsma (1983) believed that functional efficiency could be applied in Bonsmara cattle based on the assumption that selection for phenotypic traits that are influential on an animal's adaptability to the environment, will improve the animal's ability to express its reproduction and production potential.

Beef cow efficiency is significantly influenced by adaptive ability, bioregion, and size (Taylor *et al.*, 2008; Webb *et al.*, 2017). Given the inevitable negative effect that impending climate change will have on beef cattle production, it will become increasingly vital to connect genotypes to production environments. Pursuing the selection of genotypes that are better adapted to the kind of environments

likely to stem from climate change, would be an optimal way to ensure sustainable production (Scholtz *et al.*, 2013; Jordaan *et al.*, 2021).

1.2.5 Landscape genomics

Genomics is the study field whereby genetic information is identified, quantified, and analysed, using a variety of genomic tools. Genomics has grown popular in the animal genetics sector due to the potential contributions it offers to improving the understanding of biodiversity, animal production, disease susceptibility, and genetic factors underlying observed phenotypic traits (Zhang *et al.*, 2012; Kizilkaya *et al.*, 2013; Zhang *et al.*, 2013; Gurgul *et al.*, 2014). Genomics also plays a role in the assessment of livestock breeding values and the determination of genomic regions that are linked to various production traits (Hayes *et al.*, 2009; Saatchi *et al.*, 2011; Weber *et al.*, 2012; Gurgul *et al.*, 2014).

Applications of genomics for beef cattle was introduced to South Africa in 2015, by way of the state-funded Beef Genomics Program (BGP) (van Marle-Köster & Visser, 2018). The goal of the program was to establish reference populations for the 16 local beef breeds that participated, in an effort to implement genomic selection for them (van Marle-Köster & Visser, 2018). The BGP enabled Bonsmara breeders to implement genomic enhanced breeding values into their selection programs, ultimately accelerating and increasing the accuracy of genetic progress within the breed (van Marle-Köster *et al.*, 2021).

The most popular tool in livestock genomic studies is the SNP genotyping array. Commercial SNP arrays are quick, reliable, and relatively inexpensive, all while providing genotypic information on large quantities of SNPs – which are a primary source of genetic variation (Matukumalli *et al.*, 2009; Gurgul *et al.*, 2014). SNPs have become popular due to their genetic stability, responsiveness to high through-put automated analysis, and abundance in the genome (Vignal *et al.*, 2002; Yadav *et al.*, 2017). SNP genotyping arrays enable scientists to detect mutations linked to specific traits or diseases in animals. The SNP arrays are designed to describe the genetic variation of a genome of interest in the most informative way possible (Gurgul *et al.*, 2014). The first of their kind, the BovineSNP50 genotyping array (Illumina Inc., San Diego, CA) became available in 2007 and featured 54001 informative SNP probes (Goddard & Hayes, 2011; Qwabe *et al.*, 2013). Over the years, new SNP arrays have been developed and improved. Table 1.3 lists the commonly used SNP arrays in the South African beef breeding industry. Some have a greater SNP density than others and some are genus specific (*Bos taurus*/*Bos indicus*). These SNP arrays are often used in genome-wide association studies (GWAS) to identify genomic regions that contribute to natural variation in economically important traits (Goddard & Hayes, 2011).

Table 1.3 Summary of SNP array information (compiled from Reding, 2020; Visser *et al.*, 2023).

LD Panel and label	Number of SNP variants per array	Manufacturer
GeneSeek® Genomic Profiler™ 80k	76 883	GeneSeek®
GeneSeek® Genomic Profiler™ 150k	139,480	GeneSeek®
Bovine3k BeadChip	2 900	Illumina, Inc. San Diego
BovineLD BeadChip	6 909	Illumina, Inc. San Diego
BovineLD.v2 BeadChip	7 931	Illumina, Inc. San Diego
BovineSNP50 version 1	54, 001	Illumina, Inc. San Diego
BovineSNP50-24 version 3	53, 218	Illumina, Inc. San Diego
Bovine HD BeadChip	777 962	Illumina, Inc. San Diego
International Beef and Dairy, (IDB) version 3	53,450	International Dairy and Beef
Weatherbys Scientific VersaSNP 50K™	49 778	Weatherbys Scientific
Unistel-SA Stud Book 50K version 1	54,394	SA Stud Book

Landscape genomics stem from landscape genetics, which offers a unique collection of spatial analysis methods that aim to investigate the manner by which landscape variables influence genetic population structures (Storfer *et al.*, 2018). The transition to genomics occurred when SNP markers (hundreds of thousands to millions of loci) replaced microsatellite markers (hundreds to thousands of loci) as the form of genetic information visualize in studies of spatial genetic variation (Joost *et al.*, 2007; Storfer *et al.*, 2018). Landscape genomics studies commonly aim to describe spatial patterns of selection and adaptation, whereas landscape genetics studies prioritise explaining the influence of landscape variables on gene flow (Storfer *et al.*, 2018).

Spatial data for most of the inhabited regions of the globe are available due to increased and improved geographic information systems (GIS) and mapping technologies (Storfer *et al.*, 2018). Additionally, from a livestock perspective, developments in next-generation sequencing have enabled the study of the genomic basis of local adaptation for nearly any organism (Storfer *et al.*, 2018). The combination of these two factors creates a unique place for landscape genomics in the livestock industry. This analytical tool is desirable for use within the livestock industry because extensive phenotypic/genotypic records are not needed for analysis. The only requirement for obtaining information about the genome relative to the environment, is that there are molecular markers spread throughout the genome (Mdladla *et al.*, 2017). Further, substantial geographic data is also not needed as geographical coordinates can be used as proxies (Mdladla *et al.*, 2017). The information that landscape genomic analysis can offer could provide valuable insight into an animal's (or in this case the Bonsmara breed's) adaptive genetic mechanisms that influence GxE interactions (Vajana, 2017). Flori *et al.* (2019) reported that Genomic

Environmental Association (GEA) analyses assisted in identifying genetic variants associated with environmental variables, which describe adaptation to specific environments.

Application of landscape genomics to the livestock industry

Some of the limited available studies for landscape genomics in the livestock industry are listed in Table 1.4. The studies all investigated local adaptation using a landscape genomics approach and provided evidence for landscape genomics being a useful tool in detecting candidate loci associated with local genomic adaptation.

One of the primary motivations for using landscape genomic analyses in the livestock sector is that it provides an opportunity to use genomic and environmental information to detect signatures of adaptive genetic variation that could explain livestock adaptation to the specific environment. Such information can supplement typical genetic evaluations for breed improvement (Long, 2008; Rellstab *et al.*, 2015; Grummer *et al.*, 2019). Additionally, by integrating global warming models with landscape genomics analysis, it becomes possible to predict the effect of climate changes on breed survivability (Joost & Negrini, 2010).

Table 1.4 List of previously conducted landscape genomic studies on the local adaptation of various livestock species.

Species	Breed	Country	Reference
Indigenous goats	Boer, Kalahari Red and Savanna, local ecotypes, and Tankwa	South Africa	Mdladla *et al.*, 2018
Indigenous cattle	Local ecotypes	Uganda	Vajana *et al.*, 2018
Indigenous goats	34 local ecotypes	Italy	Cortellari *et al.*, 2021
Indigenous cattle	Local ecotypes	Eritrea	Goitom *et al.*, 2021
Sheep and goats	26 French sheep breeds, 16 Italian sheep breeds, and 9 Spanish sheep breeds. 8 French goat breeds, 18 Italian goat breeds, 6 Spanish goat breeds.	Mediterranean	Serranito *et al.*, 2021
Indigenous sheep	13 local ecotypes	Ethiopia	Wiener *et al.*, 2021
Indigenous cattle	11 local ecotypes	India	Bhardwaj *et al.*, 2023

The landscape genomics study conducted by Mdladla *et al.* (2018) identified 195 associated candidate genes. The authors visualize two landscape genomics approaches, LFMM and SAM. 55 genes were identified through the LFMM approach, and 140 genes were identified through the SAM approach. The authors noted *ADRA1D*, *BRAF*, *CALCRL*, *CALD1*, *EDNRA*, *ITPR2*, *PLCB1*, and *PRKG1*, to be candidate genes for local adaptation of indigenous South African goats (Mdladla *et al.*, 2018). The study investigating local adaptation of indigenous Italian goats visualize both SAM and LFMM landscape

genomics approaches (Cortellari *et al.*, 2021). The SAM approach identified 62 genes, of which *DLG1*, *HADC9*, and *KLF12* were associated with meat and growth; *BTRC*, *DENND1A*, and *PRKD1* were associated with fertility; *BTLA* was associated with rheumatoid arthritis; and *EYA3*, *KCNJ1*, and *MAPK9* were associated with circadian rhythm (Cortellari *et al.*, 2021). The LFMM approach identified two candidate genes from the four SNPs associated with the tested environmental variables. These genes were *NBEA* and *RHOBTB1* (Cortellari *et al.*, 2021).

Serranito *et al.* (2021) visualize an LFMM approach to identify 42 genes that were believed to be found in regions under selection. The authors reported that *ADAMTS20*, *DPH6*, *NBEA*, *SOX2*, *TRPC4*, *TRPC6*, *UBE2R2/UBAP2*, and *UHRF1BP1/C23H6ORF106* were associated with local adaptation of indigenous Mediterranean sheep and goat breeds (Serranito *et al.*, 2021). Local adaptation of indigenous Ethiopian sheep was investigated using a Bayesian approach by Wiener *et al.* (2021). 56 candidate genes were identified and *ARMC3*, *COL6A3*, *FHAD1*, *PLCB1*, *PRDM16*, *RXFP2*, and *SDK1* were reported to be specifically associated with adaptation to various environmental variables (Wiener *et al.*, 2021).

Vajana *et al.* (2018) visualize a SAM approach and identified a total of 42 candidate genes from the variables tested. The authors noted *PRKG1* and *SLA2* as prominent candidate genes for local adaptation of East Coast Fever (ECF) in Ugandan cattle (Vajana *et al.*, 2018). The study by Goitom *et al.* (2021) identified 1061 genes from the landscape genomic approach the authors implemented. Eight of those genes were found in genomic regions under positive selection when selection signatures were scanned. The genes were *AHSG*, *AIRE*, *ATP1B3*, *CASR*, *IFNAR2*, *PARS2*, *ROBO2*, and *SCHIP1* These genes were reported to be associated with stress and defence response, autoimmunity regulation, energy-related biosynthesis, cellular mineral homeostasis, blood coagulation, amino-acylation, morphogenesis of cells, locomotion, facial morphology, and skeletal and muscle system development (Goitom *et al.*, 2021). Bhardwaj *et al.* (2023) visualize a SAM approach for their landscape genomic analysis of indigenous Indian cattle. There were 1305 significant SNPs identified from the analysis, from which a number of genes were identified that were associated with various forms of adaptation (acclimation, adipose tissue, coat colour, cold adaptation, disease resistance, growth, light stress, meat quality, milk production, and reproduction) (Bhardwaj *et al.*, 2023).

Genome-wide association study (GWAS) and selection signatures

Genome-wide association studies (GWAS) are genotypic studies that aim to identify genetic variants that are significantly associated with variation in specific traits (Raza *et al.*, 2020). In addition to genotypic information, phenotypic information for the genetic variant being explored is required for the GWAS to determine a correlation between the trait and the variant (Hirschhorn & Daly, 2005; Bush & Moore, 2012; Santiago *et al.*, 2017). The information provided by a GWAS can be useful in guiding selection for complex traits in populations (Bolormaa *et al.*, 2014; de Vos, 2018). The principal

assumption supporting GWAS is that significant associations occur when the SNPs and causative mutations for a trait are in linkage disequilibrium (LD) (Hayes & Goddard, 2010).

Genomic selection (GS) differs from marker assisted selection (MAS) in that it considers the effect of large numbers of genes on a single trait, instead of the effect of only one or a few genes (Hayes & Goddard, 2010). This expansive exploration of gene influence on trait expression allows for potentially improved understanding of the variation observed through the expression of complex traits (Korte & Farlow, 2013; de Vos, 2018). The sample size and density of the SNP array determine the number of SNPs detected through GWAS. The larger the sample size and the higher the density of the SNP array, the greater the number of SNPs detected (de Vos, 2018). However, the success of a GWAS depends on a number of factors, including LD. The extent of LD across the genome varies between breeds and populations, therefore it is important to take into consideration.

Generally, GWAS results are visualized through a Manhattan plot, allowing for easier interpretation. With this approach however, significance levels for the associations, confidence intervals and population parameters need to be accounted for (de Vos, 2018). The significance level is displayed on the Manhattan plot as a Bonferroni corrected significance level and is usually set to 5%. In which case, p-values >0.05 signify that a trait is significantly associated with those SNPs. The confidence interval is generally set to 95%. It is important to consider population structure to avoid exaggerated associations or high rates of false positives (Hirschhorn & Daly, 2005; de Vos, 2018).

GWAS has been used exhaustively to explore growth traits in beef cattle (Berry *et al.*, 2017; Raza *et al.*, 2020) and numerous SNPs have been identified on various chromosomes for these traits. A Canadian study found four SNPs that were significantly associated with BW, pre-weaning daily gain (PDG), WW, and YW (Akanno *et al.*, 2018). These SNPs were detected on bovine chromosomes BTA1, BTA3, BTA4, and BTA21, and were associated with four candidate genes: *U6atac* (BW), *AGBL4* (WW), *bta-mir-2888-1* (PDG), and *REPIN1* (YW), respectively (Akanno *et al.*, 2018). Makina *et al.* (2015) determined several genes to be associated with growth, metabolic processes, muscle organ development, and skeletal development. The growth and metabolic process associated genes were *AJAPI* (BTA7), *DDX19A*, *IGFBP4*, *KCNB1* (BTA8), *MYO6*, and *TGFB1*. The genes associated with skeletal and muscle organ development were *EFHD2*, *KIAA1797* (BTA6), *MTPN* (BTA6), and *TMEM51* (Makina *et al.*, 2015).

Additionally, another study (Buzanskas *et al.*, 2014) found BW, WW, and YW to be associated with SNPs located close to close to genes on various chromosomes. BW was significantly associated with SNPs on BTA4 (close to *DPP6*) and BTA9 (close to *MANEA* and *LOC783932*). WW was significantly associated with SNPs found on BTA 6 (near *FARSB*) and BTA11 (near *RALGDS*). YW was significantly associated with SNPs located BTA7 (near *ALDH7A1*, *C7H5orf48*, *LMNB1*, *LOC100848523*, *MARCH3*, *MIR2458*, and *PHAX*); BTA22 (near *CDCP1*, *CLEC3B*, *EXOSC7*, *LARS2*,

LOC614114, *LOC101907967*, *LOC101908013*, *LOC101908094*, *LOC101901958*, *TMEM158*, and *ZDHHC3*); and BTA27 (close to *LOC101904868*) (Buzanskas *et al.*, 2014). The use of GWAS for growth trait investigation has provided a greater understanding of the genetic variation expressed with these traits. It has also provided useful insight into the candidate genes that could be targeted in growth trait selection.

Selection signatures

Selection leads to alterations of regions of the genome that result in selection signatures. Identifying selection signatures can facilitate the detection of genomic regions that have been targeted by selection due to being (or having been) functionally important (Makina *et al.*, 2015). The relevance of identifying selection signatures lies in their contribution to understanding mechanisms associated with traits that have been exposed to intensive selection. Moreover, selection signatures facilitate the successful annotation of significant functional genomic regions (Makina *et al.*, 2015). However, one of the primary challenges to identifying selection signatures in livestock is the existence of SNP ascertainment bias. The SNP assays utilised for analysis in most studies, contain common SNPs – thus, any information regarding levels of LD or the distribution and variability of allele frequencies, would be heavily influenced by those common SNPs (Makina *et al.*, 2015).

Adaptation to specific environments has been observed to leave unique signatures in the genome. This is due to an abundance of desirable allele frequencies or frequencies of neutral markers that are in LD with favourable alleles (Smith & Haigh, 1974; Ben-Jemaa *et al.*, 2020). These selection signatures may expose genes that are connected to traits under selection. Selection signatures can be identified by statistical methods that are broadly classified as intra-population statistics and inter-population statistics (Saravanan *et al.*, 2020). Intra-population statistics compare the genomic data within populations to identify selection signatures. Example approaches include site frequency spectrum, LD, and reduced local variability (i.e., runs of homozygosity (ROH)) (Weigand & Leese, 2018; Saravanan *et al.*, 2020). Inter-population statistics are dependent on methods that take the extent of differentiation caused by locus-specific allele frequencies between populations into account (Zhao *et al.*, 2015; Saravanan *et al.*, 2020). These approaches include single site differentiation (i.e., fixation index (F_{st})) and haplotype-based differentiation (i.e., haplotype-based extension of FLK (hapFLK)) (Fariello *et al.*, 2013; Saravanan *et al.*, 2020).

Environmental adaptation is comprised of complex traits that are influenced by many genes and metabolic pathways. Thus, it can be challenging to identify only a few loci under positive selection (Kemper *et al.*, 2014; Freitas *et al.*, 2021). Despite this challenge, scientists have been able to identify candidate genomic regions associated with adaptability traits in various cattle breeds by studying selection signatures. One such study was conducted by Freitas *et al.* (2021), who isolated a few candidate genes related to adaptability traits by use of intra- (ROH) and inter-population (Fst, and

HapFLK) statistical analysis methods. The authors analysed 32 worldwide cattle breeds (with specific focus on Asian cattle) for genomic regions that were potentially under selection for heat tolerance. The statistical analysis methods revealed that there were candidate genes associated with physiological pathways and processes such as feed intake, heat-shock proteins, and metabolic activity (Freitas *et al.*, 2021). Additionally, a similar study was conducted by Ben-Jemaa *et al.* (2020), where genome-wide scans for selection signatures provided insight into the adaptive ability of North African cattle. Outlier genomic windows were detected, wherein candidate genes associated with adaptation to drought, forage scarcity, and infectious diseases were located (Ben-Jemaa *et al.*, 2020).

Studies such as those conducted by Ben-Jemaa *et al.* (2020) and Freitas *et al.* (2021) can be pivotal in identifying biological factors that influence adaptation and physiology, and subsequently production, in livestock. The genomic regions that were detected by the studies could potentially provide insight to the locations of candidate loci involved with environmental adaptation. Such information can be beneficial to future landscape genomic studies that seek to investigate the influence of the environment on specific SNP variants instead of taking a genome-wide association approach (Gurgul *et al.*, 2014). Detail on selection signatures and its application was beyond the scope of this study.

1.2.6 The methodology of landscape genomics analysis

Landscape genomic approaches generally require three forms of data for analysis: geographic coordinate data, environmental variable data, and genetic information (Joost *et al.*, 2007). The environmental data typically included in analysis include rainfall, temperature, relative humidity, vegetation and topography (Mdladla *et al.*, 2018), while the genetic information consists of SNP markers (Storfer *et al.*, 2018). When genetic information is not readily available and must be sequenced specifically for a landscape genomic study; one of the popular sequencing methods visualize is restriction-associated digest DNA sequencing (RAD-seq.) (Andrews *et al.*, 2016; Lowry *et al.*, 2017). This approach is appealing to scientists because, although reference genomes exist for livestock species, a reference genome is not required for RAD-seq. execution. It also enables the identification of thousands to millions of SNPs (Storfer *et al.*, 2018; Georges *et al.*, 2019).

Genome scans can be used to test for loci that are under selection. Two approaches are considered for loci detection: (a) differentiation outlier methods (previously F_{ST}-outlier tests), and (b) GEA tests (Rellstab *et al.*, 2015; Storfer *et al.*, 2018). Of interest to this study are the GEA tests, as they are designed to identify significant associations between allele frequencies and environmental variable variation (Rellstab *et al.*, 2015). Table 1.5 summarises the compulsory information required for computational landscape genomic association tests and the optional components that can be applied.

Table 1.5 Necessary and optional information for livestock landscape genomic analyses and expected outputs (adapted from Joost & Negrini, 2010).

	Input	**Output**
Necessary information required for calculation of association models	Environmental data specific to the bioregion under investigation (topographic and/or climatic information, descriptive information regarding the production system i.e., pasture type, average annual precipitation, soil). Genetic data specific to the bioregion under investigation (SNPs).	Working theories regarding the roles of the genomic regions associated with the genetic markers. Potential habitat maps for relevant the breeds/populations. Forecasts of potential effects of climate change. Forecasts of potential effects of landscape change resulting from human activity.
Optional extra information on diseases and farming systems	Disease information specific to the bioregion under investigation. Farming system data specific to the bioregion under investigation (socio-economic information pertaining to farms, management systems, objectives, resources).	Working theories about relationships between specific genomic regions and disease occurrence.

Several statistical model approaches have been developed into computer programs that facilitate the detection of GEAs, since landscape genomics was first introduced in the early 2000s (Rellstab *et al.*, 2015; Li *et al.*, 2017). A few of those models and programs are the spatial analysis method (SAM) (Joost *et al*, 2007), the Bayesian method (BAYENV) (Coop *et al.*, 2010), and the latent factor mixed model method (LFMM) (Frichot *et al.*, 2013). Details on these models are given in Table 1.6. When compared to other GEA test models, LFMMs were reported to be more effective models for detecting loci under selection, when applied to individual-based sampling designs (Lotterhos & Whitlock, 2015; Forester *et al.*, 2018; Gugger *et al.*, 2021).

Table 1.6 Summary of the different statistical models/computational programs used for landscape genomics (Frichot & Francois, 2015a; Rellstab *et al.*, 2015).

Model name	**Model type**	**Advantages**	**Disadvantages**	**Software/*R* package**
Spatial analysis method (SAM)	Linear regression model	Can distinguish slight differences in allele frequencies between populations separated by location and landscape.	Does not correct for confounding effects. Requires pre- and post-treatments: • Population structure analysis (pre)	*SAM* (Joost *et al.*, 2007); *SAMβADA* (Stucki *et al.* 2017)

			• Control of false discovery rates (post) • Visualisation of results (post)	
Bayesian method (BAYENV)	Linear regression model	Can differentiate slight variations in allele frequencies between populations separated by location and landscape. Corrects for confounding effects (genomic background, isolation-by-distance patterns, population structure).	Requires pre- and post-treatments: • Population structure analysis (pre) • Control of false discovery rates (post) • Visualisation of results (post)	*BAYENV* (Coop *et al.*, 2010); *BAYENV2* (Günther & Coop, 2013)
Latent factor mixed model method (LFMM)	Linear mixed model	Can differentiate slight variations in allele frequencies between populations separated by landscape and location. Corrects for confounding effects (genomic background, isolation-by-distance patterns, population structure). Enables separation of neutral genetic variation from adaptive genetic variation to identify adaptive genomic selection signatures.	Requires pre- and post-treatments: • Population structure analysis (pre) • Control of false discovery rates (post) • Visualisation of results (post)	*LFMM* (Frichot *et al.*, 2013); *LEA* (Frichot & Francois, 2015)

Methodology of LFMM approach

Generally, there are two predominant steps involved in GEA studies (Frichot & Francois, 2015a; Lotterhos & Whitlock, 2015). The first involves the evaluation of population genetic structure from provided genomic data, which is followed by assessing the various factors that may have an influential on the interpretation of results. The second step involves correcting any biases attributable to population structure results from step one, as well as any other confounding factors (Frichot & Francois, 2015a). LFMM methods test associations between genomic information and ecological variables through a linear regression model, where the allele frequencies of the animals are stored in a genotypic matrix and the ecological variables are factors of the genotypic matrix. The model

$$G_{il} = \mu_l + \beta_l^T X_i + U_i^T V_l + \epsilon_{il} \qquad \text{Eqn. 1.1}$$

consists of a locus-specific effect, μ_l; a *d*-dimensional vector of regression coefficients, β_l; *K* latent factors, U_i; their corresponding [*K*] values, V_l.; where, *i* denotes an individual and *l* denotes a locus; and Gaussian variables with variance σ^2 and a mean of zero, ϵ_{il} (Frichot *et al.*, 2013). The associations are tested by estimation of unobserved latent factors that model the confounding effects. In essence, the latent factors consist of layers of population structure from shared, background genetic variation (Frichot & Francois, 2015a). Once confounding effects have been corrected for, any notable associations observed between allele frequencies and specific ecological variables, are taken as indications of selection at the relevant loci (Frichot *et al.*, 2013).

In the instance that the investigated ecological variables do not influence genetic variation, LFMM are assumed to produce uniformly distributed P-values, given correct calibration of the tests (Frichot & Francois, 2015a). To test whether a specific model is calibrated correctly, LFMM can be run with unique numbers of latent factors. Optimal LFMM association test performance has been observed for latent factor values that are similar to the number of significant components or clusters in a principal component analysis (PCA) (Frichot *et al.*, 2013). The output of the LFMM association tests is also noteworthy. Results consist of P-values and Z-scores; by combining the Z-scores from multiple runs and appropriately controlling the false discovery rate, lists of candidate loci can be generated (Benjamini & Hochberg, 1995). The candidate loci can be used to identify genes for annotation, to determine SNPs that may be influenced by the environment.

Challenges to the LFMM approach

It has been reported that one of the primary challenges to LFMM methods is deciding on an appropriate K value (Lotterhos & Whitlock, 2015). Choosing incorrect K values can skew the output and be detrimental to the interpretation of results. K values that are too large (K>50) result in more conservative tests and a decline in the power to reject neutrality (Patterson *et al.*, 2006). Therefore, although challenging, finding the optimal K value is key. A suitable K value can be equal to a reasonable estimate of the number of genetic clusters, determined from clustering programs such as STRUCTURE or ADMIXTURE (Patterson *et al.*, 2006; Frichot *et al.*, 2013). Observations have shown that when the number of latent factors (K) are equal to the number of genetic clusters (observed in the PCA), LFMM tests perform optimally (Frichot *et al.*, 2013).

Another limitation to this model type is that it requires pre- and post-treatments of data and output (Frichot & Francois, 2015a). Prior to the LFMM analysis, analysis of population structure must be conducted. Post LFMM analysis, the false discovery rates need to be controlled for and the results need to be visualized in plots and graphs (Frichot & Francois, 2015a). Despite these challenges, it has been reported that when calibrated correctly and when the K value was true, LFMM produced lower false negative and false positive rates than other models (Lotterhos & Whitlock, 2015).

Landscape Ecological Association (LEA) analysis

Landscape ecology association (LEA) analysis refers to an Rstudio package developed by Frichot and Francois (2015a, 2015b). The LEA v.3.8.0 package is an integrated framework that combines the two main steps of a GEA: population structure genetic analysis and ecological association study (Frichot & Francois, 2015a). The package allows for population structure analysis by way of principal component analysis (PCA) or alternatively, non-negative matrix factorization algorithms (sNMF) (Frichot *et al.*, 2013; Frichot *et al.*, 2014). The ecological association of allele frequencies with ecological gradients is conducted through LFMM. LEA v.3.8.0 also comprises methods for statistical model calibration and false discovery rate control (Frichot & Francois, 2015a). Population structure analysis allows for the assessment of any genetic factors that may influence the interpretation of results. The ecological association study includes correction for biases caused by population structure and other confounding factors (Frichot & Francois, 2015a). The functions implemented in LEA can process very large genomic datasets without compromising the R program memory. Thus, by performing these two required analytical steps within a unified framework, algorithmic speed and memory allocation are visualized, whilst the statistical analysis flexibility and the visualization methods of Rstudio are conserved and benefited from (Frichot & Francois, 2015a).

Prerequisites for using a landscape genomics approach

The success of landscape genomics analysis is dependent on adequate regional and international genotypic data and sampling, which may be lacking in some studies. The costs associated with regional sampling and genotyping may hinder developing countries, such as South Africa, from conducting landscape genomic studies (Mdladla *et al.*, 2017). Thus, more international studies and pooling of resources from international sources is required for the overall success of landscape genomics (Mdladla *et al.*, 2017).

1.2.7 Conclusion

The environment plays an important role in beef cattle growth performance, given that extensive farming systems are predominantly implemented in the South African beef industry. The full extent of the environmental influence remains unclear as mature Bonsmara cow sizes vary between geographical regions and thus, standard points of reference for maintenance requirements or adaptive abilities do not exist. This creates challenges when breeders want to select for improved growth performances amidst rising levels of climate change. Landscape genomics offers opportunities to investigate associations between animal genotypes and the environments they are exposed to. The results from landscape genomics analyses should identify potential candidate genes that may aid in genomic selection for improved growth performance and adaptability to environment.

Chapter 2. Materials and Methods

2.1 Introduction

For this study, 4679 Bonsmara genotypes generated within the Bovine Genome Project (BGP) and Red Meat Research & Development (RMRD) projects, as well as through private breeders, were made available for the study with the required consent of the SA Bonsmara Breeders Society. Ethical approval for use of the external data was granted by the Ethics Committee from the University of Pretoria (NAS274/2020).

2.2 Materials

2.2.1 Genotypes

A total of 4679 genotypes were obtained from SA Stud Book representing nine different geographical regions in South Africa. Only animals that were born in the same province in which they were raised, were included in the study. Most animals that complied with this criterion, were located in the Eastern Cape, Free State, and North-West provinces. Only these three geographical regions were included in the current project, as there were too few animals in the other provinces to generate statistically significant results. All male animals were removed from the dataset as most bulls moved between provinces from birth to maturity when sold. The editing process is described in Figure 2.1.

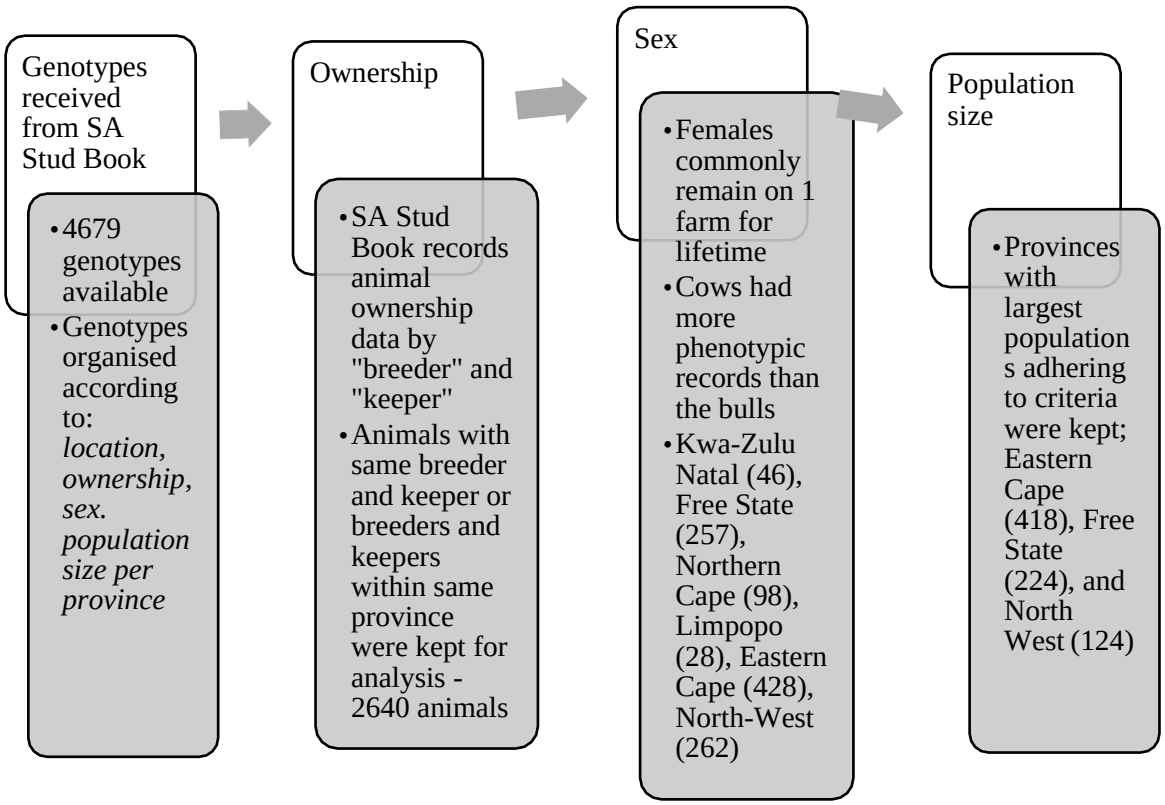


Figure. 2.1 Diagram breaking down the editing and pruning of raw data.

To be eligible for analysis, animals were required to have been born and raised at a single geographical location. The selected female animals had been born and raised on the same farm and phenotypic recordings for the traits of interest were available. After editing and pruning, 766 genotyped cows representing the Eastern Cape, Free State and North-West provinces were included in the analysis (Table 2.1).

Table 2.1 Bonsmara genotypes representing different regions of South Africa included in this study.

Province	Number of breeders	Number of genotypes
Eastern cape	7	418
Free state	15	224
North-west	20	124
Total	42	766

The SNP arrays used to genotype all the cattle are listed in Table 2.2. The second column indicates how many cattle, total, are recorded on each panel, and the proportion of bulls and cows.

Table 2.2 Complete genotype dataset obtained from SA Stud Book.

LD Panel and label	Number of cattle per array	Number of SNP variants per array
GeneSeek® Genomic Profiler™ 80k	597 cattle: 195 cow, 402 bulls	76 883
GeneSeek® Genomic Profiler™ 150k	1952 cattle: 1123 cows, 829 bulls	139 480
International Beef and Dairy, (IDB) version 3	829 cattle: 474 cows; 355 bulls	53 450
Weatherbys Scientific VersaSNP 50K™	1525 cattle: 399 cows, 1126 bulls	49 778

To visually represent the geographic information of the landscapes being investigated, a dataset of the coordinates for farms from the respective provinces was uploaded to an online interactive data visualization software Tableau version 2021.2.7 (Tableau Software, LLC, Seattle, WA). This software allows for data to be represented through maps, graphs, charts, and diagrams (Hamersky, 2016). Figure 2.2 was produced to illustrate the spread of farms in the three provinces. The blue, red and black boxes refer to the North-West, Free State and Eastern Cape provinces respectively. The labels indicate respective weather stations, where the environmental data were obtained from, and their coordinates.

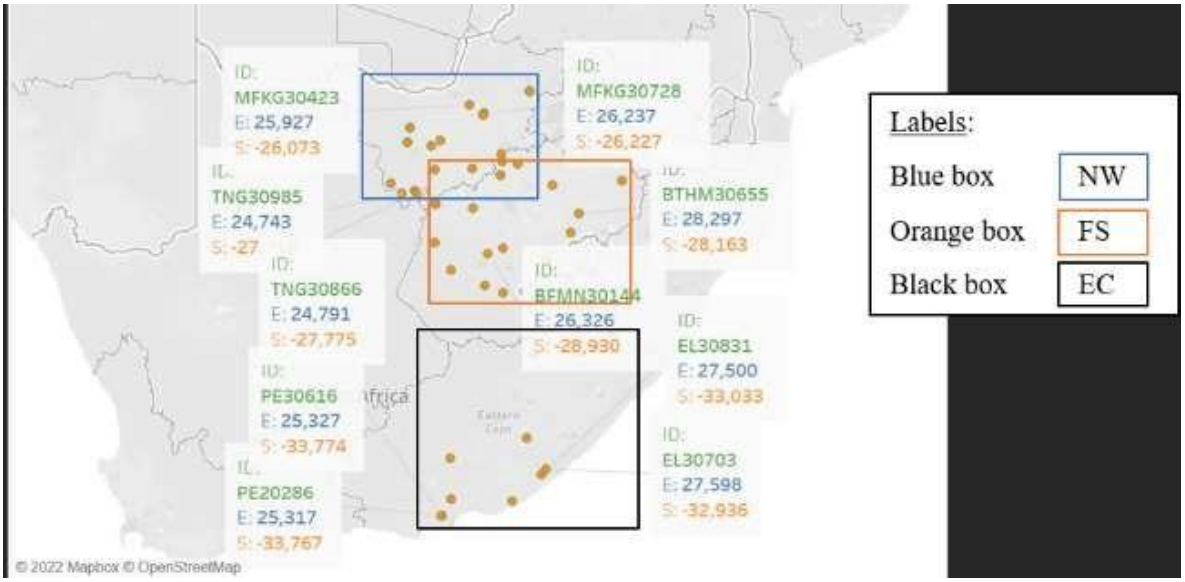


Figure 2.2 Map indicating geographical coordinate locations of animals on farms in North-West, Free State and Eastern Cape (Tableau, 2021).

2.2.2. Phenotypes

The production traits that were taken into consideration to investigate the effect of adaptation on growth performance consisted of birth weight (BW), weaning weight (WW), yearling weight (YW) and 18-month weight (18MW). Measurement data for these traits was available for most of the cows used in analysis, the number of available phenotypes per province are listed in Table 2.3. These phenotypes were utilised in a GWAS to identify any existing significant SNPs associated with growth performance, specifically 18MW.

Table 2.3 Number of phenotypes available per trait, per province.

	Number of animals	**BW**	**WW**	**YW**	**18MW**
Eastern Cape	418	418	400	332	314
Free State	224	215	211	163	122
North-West	124	123	116	102	89
TOTAL	766	756	727	597	525

The average phenotypic performances for the traits are given in Table 2.4. The average performance measurements were calculated from the data obtained from SA Stud Book.

Table 2.4 Age of measurement and average performance for weight traits in cows per province.

Trait	**BW**	**WW**	**YW**	**18-MW**
Age of measurement	Within 3 days of birth	151-270 days	271-450 days	451-634 days
Eastern Cape	35.12kg	232.8kg	272.32kg	374.79kg
Free State	34.43kg	225.35kg	276.69kg	357.37kg
North-West	33.62kg	224.60kg	262.75kg	335.05kg
Average overall	35.2kg	228.0kg	271.16kg	356.27kg

2.2.3 Environmental variables

Environmental data from the Eastern Cape, Free State and North-West provinces was requested from the SA Weather Bureau. They provided raw data consisting of daily recordings for maximum and minimum temperatures, maximum and minimum relative humidity and precipitation from two weather stations in each province. The recordings ranged from 2001 to 2021, however, there were many inconsistencies in the measurements with some years missing entire months of recordings, especially the years 2001-2015. The recordings from the years 2016 to 2021 were utilised for analysis. The farm locations for all the animals were compared to the coordinates of the two weather stations in each province. The values corresponding to the closest weather station were the ones applied to the relevant farms.

The maximum and minimum relative humidity data did not vary significantly, thus the recordings were consolidated and placed under the single variable of "relative humidity" or "Rhum", to simplify analysis and minimize run time. The data was processed to obtain an average annual value for each variable, which corresponded to the respective weather stations. This was as per recommendations by the authors of the Landscape Ecology Association (LEA) v.3.8.0 software (Frichot & Francois, 2015a) and observations of environmental input file formats of sample data sets provided by the authors as public access (Frichot & Francois, 2015b, Mdladla *et al.*, 2018). Table 2.5 shows the average measurements for each variable over the 2016 to 2021 period, calculated from the weather data obtained from the SA Weather Bureau in 2021. The values in Table 2.5 were those used in the environmental data sets for the LEA analysis.

Table 2.5 Climatic variable data for provinces (averages from 2016 to 2021).

Weather station location	**[a]Compno**	**Summer month temperature** (°C)	**Winter month temperature** (°C)	**Relative humidity (%)**	**Average precipitation (mm)**
PE (Eastern Cape)	30616	28.55	7.12	63.82	445.85
East London (Eastern Cape)	30703	25.65	10.6	65.46	418.42
Taung (North-West)	30985	31.3	3.73	53.04	488.03
Mafikeng (North-West)	30728	28.27	5.9	53.97	477.81
Bethlehem (Free State)	30655	26.52	2.02	60.46	654.94
Bloemfontein (Free State)	30144	30.52	3.64	49.71	558.81

[a]Compno, weather station reference number

2.3 Statistical analyses

2.3.1 Quality control of genotypic information

Genotypic information for the cattle was obtained from SA Stud Book and originated from the following SNP array panels; GGP80k, GGP150k, IDB and Versa 50k. Quality control (QC) was conducted using Plink v1.90b6.18 64-bit (Purcell *et al.*, 2007). QC was executed for each panel, for sample call rate, SNP call rate, minor allele frequencies and Hardy-Weinberg Equilibrium (HWE) with cut-off levels being 0.05. The cut-off level for the HWE was 0.0001. These cut-off levels allowed for the removal of the cattle that had more than 5% missing genotypes, and the SNPs that were missing in more than 5% of the cattle, had a MAF lower than 5%, and violated the HWE ($p<0.0001$). The common SNPs between the arrays were then extracted and the data sets were merged. After QC, 766 cows remained, and 25 272 common SNPs were extracted from the panels and used for downstream analysis.

Principal component analysis (PCA) plots were created for the cows from each province using the GCTA64 v.1.93.2 (Yang *et al.*, 2011) program and Microsoft Excel 365 v.2022 (Build 14931.20132) (Microsoft Corporation, 2018). These plots were created to compare the genomic relationships between Bonsmara populations in the respective provinces. Before the plots were constructed, genetic relationship matrices were calculated from the 25 272 common SNPs using GCTA64 v.1.93.2 (Yang *et al.*, 2011). Eigenvalues and eigenvectors were subsequently generated for the principal components. The PCA plots were produced through Microsoft Excel 365 v.2022 (Microsoft Corporation, 2018) by plotting the eigenvector values; first, principal component (PC) 1 vs. principal component (PC) 2 and second, PC 1 vs. PC 3.

From the PCA plots, it was observed that there was an outlier group of animals originating from the Eastern Cape. Referral back to the list of owners and the animals that belonged to specific breeders, revealed that all the animals in the cluster belonged to a single breeder. To determine the population substructure of these animals, their identity by descent was calculated. Plink v1.90b6.18 64-bit (Purcell *et al.*, 2007) was utilised to calculate the identical by descent (IBD) values. The average proportion of IBD was calculated for the subgroup of animals and those exhibiting the highest relatedness were identified. Of the 55 animals in the subgroup, 29 exhibited values greater than 0.04 (list of cows found in Addendum A). Despite the high relatedness of the subgroup, none of the animals were removed from the analysis data set. Due to the exploratory nature of the project a larger data set was considered optimal to conduct a more extensive investigation into the environmental influence on genetic expression.

2.3.2 Admixture

To gain a better understanding of the ancestral history of the Bonsmara populations that were used in analysis, the genomic information was analysed with the program, ADMIXTURE version 1.3 (Alexander *et al.*, 2009).

ADMIXTURE version 1.3 (Alexander *et al.*, 2009) required a K value, which represents the number of ancestral populations related to each individual in a herd (Alexander *et al.*, 2020). K values from 1 to 20 were tested through ADMIXTURE (Alexander et al., 2009) to obtain the necessary Q files. It was decided to use a K value that corresponded to the number of genetic clusters in the PCA plots. Therefore, the analysis investigation was continued with a K value of 5.

The Admixture was visualised in a graph created by the program, Genesis-windows10-java8-0.3.0-210808a (Buchmann & Hazelhurst, 2014). The Q file for K=5 and the .fam file of the files used for admixture, were uploaded to the Genesis program as was required to execute the command "New Admixture graph".

2.3.3 Climatic variable correlations

The amount of existing correlation between the environmental variables for each province was investigated through the *rstatix* package in RStudio (Kassambara, 2023). The Pearson correlation method was specified to calculate the coefficients (Schober *et al.*, 2018). The correlation coefficients were presented in table format. The tables were created using Microsoft Excel 365 v.2022 (Microsoft Corporation, 2018).

2.3.4 LEA Analysis

LEA v.3.8.0 (Francois & Frichot, 2015a) was used to analyse the data. The LEA v.3.8.0 R package allowed for a genome-wide test for local adaption (Francois & Frichot, 2015a). This was done through statistical tests that employed latent factor mixed models (LFMM) which fall under the GEA analysis type (Lotterhos & Whitlock, 2015). The LFMM utilise a Markov chain Monte Carlo (MCMC) algorithm to sample from a probability distribution (Mdladla *et al.*, 2018). The MCMC algorithm enabled a regression analysis, whereby the confounding variables were modelled with unobserved, latent factors (Frichot & Francois, 2015b; Mdladla *et al.*, 2018). The regression analysis and variable modelling enabled the estimation of correlations that existed between the environmental variables and allelic frequencies (Frichot & Francois, 2015b). With LEA v.3.8.0, the environmental variables were considered as fixed effects (Frichot *et al.*, 2013). Although LEA is an integrated framework that enables population structure analysis (pre-processing step) in addition to the LFMM analysis, the in-script steps for the population structure analysis were excluded as they were executed externally using GCTA64 v.1.93.2 (Yang *et al.*, 2011) and ADMIXTURE version 1.3 (Alexander *et al.*, 2009) software as explained previously in this section.

Pre-processing

The LEA analysis was conducted on an environmental file and a genotype file. Analysis was run separately for each individual environmental variable and the genotypes from the relevant province; thus, it was run 12 separate times. The environmental file was in *.env* format and the genotype (SNP)

file was in *.ped* format (Frichot & Francois, 2015b). The *.ped* file was converted into *.lfmm* and *.geno* formats; which were relevant to the remainder of the analysis. The environmental variables: summer month temperature, winter month temperature, relative humidity, and precipitation, and their respective values, were placed into individual *.env* files and analysis was run on each file separately. The *.env* files were initially created with an additional column containing the animal IDs (in numerical order) so as to keep track of which variable was linked to which animal. The animal ID column was removed for analysis as the LEA v.3.8.0 only accepted *.env* files with a single column of values.

The parameters of the run were set to the default settings as these were appropriate for the dataset sizes (Frichot & Francois, 2015b). The iteration number, the Gibbs Sampler algorithm, was set to 10 000. The burnin number was set to 5 000, which is half of the total number of cycles – as recommended (Frichot & Francois, 2015b). The number of run repetitions was set to 10 to gain increased reliability of results and reduced false positives. The K values (K=3 (EC), K=1 (FS), K=1 (NW), overall K=5) were determined from the PCA plots.

Post-processing

The LEA software required some post-processing of the output. The z-scores for multiple runs were combined using the Fisher-Stouffer method and p-values were re-adjusted to increase the power of the LFMM test statistic (Lipták, 1958; Frichot & Francois, 2015b). The re-adjusted p-values were then applied to the Benjamini-Hochberg procedure by a set of commands to obtain the list of candidate loci (Frichot *et al.*, 2013; Frichot & Francois, 2015b). LEA included the control of false discovery rates; therefore, it did not have to be executed separately as a post-processing step. The visualisation of the results was also included in the R. script provided by the LEA authors; thus, it also did not have to be executed separately.

The final post-processing step was to annotate the candidate loci. The SNP names were determined by cross-referencing the candidate loci list to the SNP *.map* file. The fixation indices (F_{ST}) of all the SNPs were calculated using PLINK and the SNP *.map* file. The 20 SNP variants with the highest F_{ST} values were identified for each respective LEA v.3.8.0 analysis candidate loci result list and were selected for gene annotation. The chromosome and base pair positions were then used to identify the SNP IDs through the European Variant Archive (EVA, http://www.ebi.ac.uk/eva). The SNP IDs were placed into the Ensembl Variant Predictor (McLaren *et al.*, 2016) to obtain the Ensembl IDs for genes associated with the variants. Lastly, the Ensembl gene IDs were placed into the Panther Classification System data base version 14 (Mi *et al.*, 2019) to obtain their relevant annotation information.

The Panther Classification System distributed the molecular functions and biological processes of the identified genes into various categories, each of which was distinguished by proportional percentage

(Mi *et al.*, 2019). The proportional percentages were visualised in proportional circles created in Microsoft Excel 365 v.2022 (Microsoft Corporation, 2018). "Circle" shapes were chosen for the various molecular function and biological process proportions. To visualise the different proportions, the radii of the circles were adjusted using the following formula:

$$r = (2 \times \sqrt{(\frac{area}{\pi})}) \div 400 \qquad \text{Eqn. 2.1}$$

The "*area*" component of the formula equated to the proportional percentage value of the molecular function or biological process, which was provided by the Panther Classification System (Mi *et al.*, 2019). The adjusted diameter was divided by 400 to obtain a decimal value which was applied to both the "*height*" and "*width*" size modifiers (found in the "Size" tab of the "Shape Format" ribbon in Microsoft Excel 365 v.2022) to adjust the proportional shape of the circle (Microsoft Corporation, 2018). To distinguish the functions/processes from one another, the circles were *filled* with different colours and keys were created to indicate which coloured circle belonged to which molecular function or biological process.

2.3.5 GWAS

A GWAS was conducted on the common 25 272 SNPs to determine if any of them were significantly associated with BW, WW, YW, or 18MW. The GWAS results were used as comparative benchmarks to the SNPs from the candidate loci identified by the landscape genomics analyses. Plink v1.90b6.18 64-bit (Purcell *et al.*, 2007) was utilised for the association studies and separate phenotype files were created for each trait. A separate association study was conducted for each trait. The Manhattan plots for each of the association studies were produced using the following packages *tidyverse*, *ggtext* and *ggplot*, in RStudio 2022.12.0 Build 353 (Purcell *et al.*, 2007; Wickham, 2016; Wickham *et al.*, 2019; RStudio, 2022; Wilke & Wiernik, 2022). The significance level for the Manhattan plots was set to $5x10^{-8}$ as this is considered the significance level of a convincing association (Pirinen, 2023).

Chapter 3. Results

The results are organised into three categories: input, output, and gene annotation. The input results comprise those obtained during the preparation of the input files; PCA plots, ADMIXTURE plot, and an environmental association plot. The output results are those derived directly from the LEA v.3.8.0 analysis; lists of detected candidate loci and built-in GWAS Manhattan plots and p-value histograms (Addendum C), for each environmental variable tested. The gene annotation results consist of those obtained from the ontological deduction process; candidate loci, SNP ID, and Ensembl gene ID identification, and gene annotation. Lastly, the GWAS analysis is displayed in a Manhattan plot with significant associations to 18MW.

3.1 Input results

Figure 3.1 is a PCA plot for PC1 vs PC2. There are three predominant clusters visible in this plot; with two being formed by individual clusters of Eastern Cape animals and the third being an amalgamation of the remaining Eastern Cape cows, all the Free State cows and all but one of the North-West cows. There appears to be an outlier from the North-West that falls to the west of the primary cluster consisting of animals from all three provinces. This animal was excluded from subsequent analyses, leaving the final number of cows for analysis at 765.

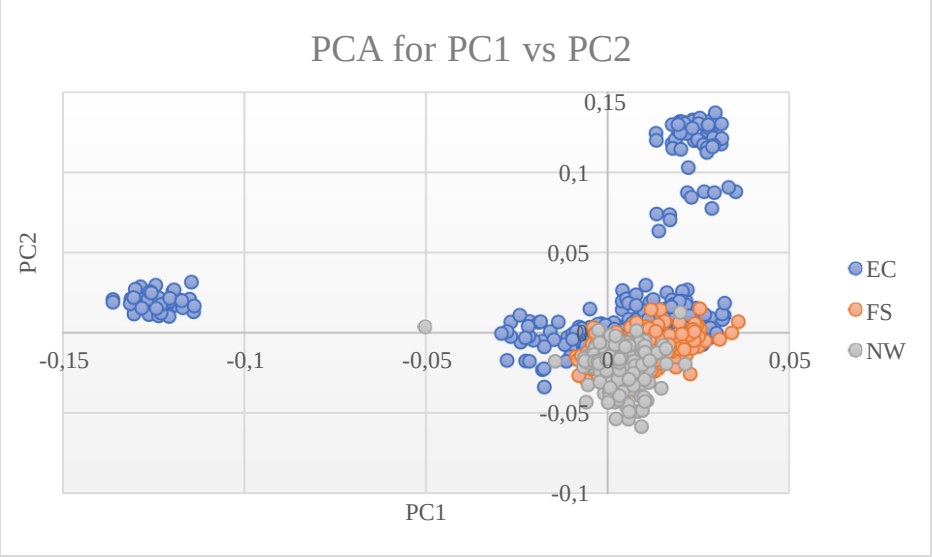


Figure 3.1 PCA of PC1 vs PC2, for the 766 Bonsmara cows included in the study. EC = Eastern Cape province, FS = Free State province, NW = North-West province.

Figure 3.2 is a PCA plot of PC1 vs PC3. In this plot, there are only two clear clusters. The primary cluster consists of animals from all three provinces, as well as the Eastern Cape cluster that fell in the top right quadrant of Figure 3.1. The second cluster was the other outlier Eastern Cape cluster that fell in the top left quadrant of Figure 3.1. There remained a single North-West outlier in the middle of the two clusters, which was excluded from the analyses.

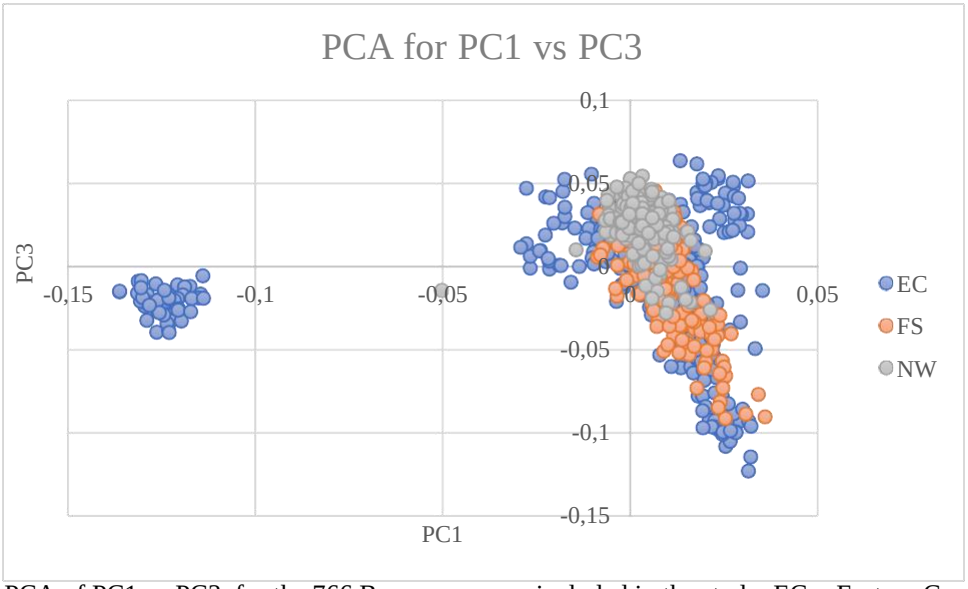


Figure 3.2 PCA of PC1 vs PC3, for the 766 Bonsmara cows included in the study. EC = Eastern Cape province, FS = Free State province, NW = North-West province.

Although most of the animals clustered together, there appeared to be even tighter clusters within province. This was especially observed within the NW cluster of animals. This was less evident with the cows from the Eastern Cape province. The 55 Eastern Cape animals that clustered in the top left quadrant of Figure 3.1 and in the bottom left quadrant of Figure 3.2 all belonged to a single farmer. The relatedness of these animals was further investigated by means of their IBD values. An even split was observed between highly related (>0.04) and moderate to lowly related (0.01 to 0.039) animals, with 28 animals falling into the highly related category and 27 animals falling into the moderate to lowly related category. The animals were not removed from analysis, but their clustering was taken into consideration and factored into the K value for the Eastern Cape LEA v.3.8.0 analysis.

The admixture plots (Figure 3.3, 3.4, 3.5) illustrated a diverse range of ancestry for all the animals. The admixture plots corresponded to the clustering observed in the PCA plots (Figures 3.1 and 3.2). Figure 3.3 illustrated the population substructure for K=3, considering only the three provinces. Figure 3.4 illustrated the population substructure for K=4, taking into consideration the top left quadrant cluster presented for the EC cows in Figure 3.1. Figure 3.5 illustrated the population substructure for K=5, additionally taking into consideration the top right quadrant cluster for the EC cows in Figure 3.1. The genetic relatedness based on breed composition was evident across the plot as was observed by the proportion of admixture indicated by the proportions of red, blue, and green in the vertical lines that represented each individual. The group of Eastern Cape cows that exhibited larger proportions of green and smaller proportions of red, pink, blue, and yellow represented those that form the two separate clusters observed in the PCA plots (Figures 3.1 and 3.2).

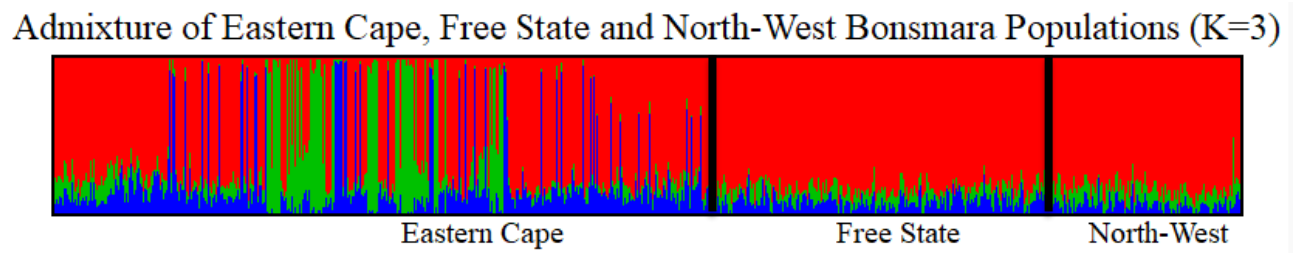


Figure 3.3 Admixture plot for Bonsmara populations from EC, FS, NW (K=3). Black lines indicate the separation of cattle by province.

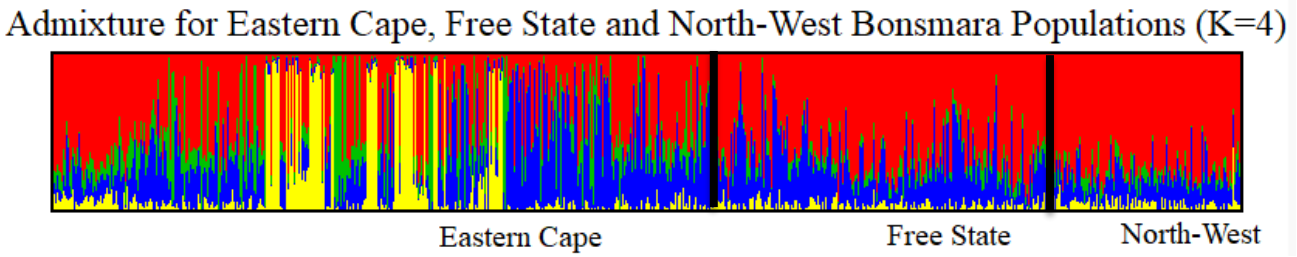


Figure 3.4 Admixture plot for Bonsmara populations from EC, FS, NW (K=4). Black lines indicate the separation of cattle by province.

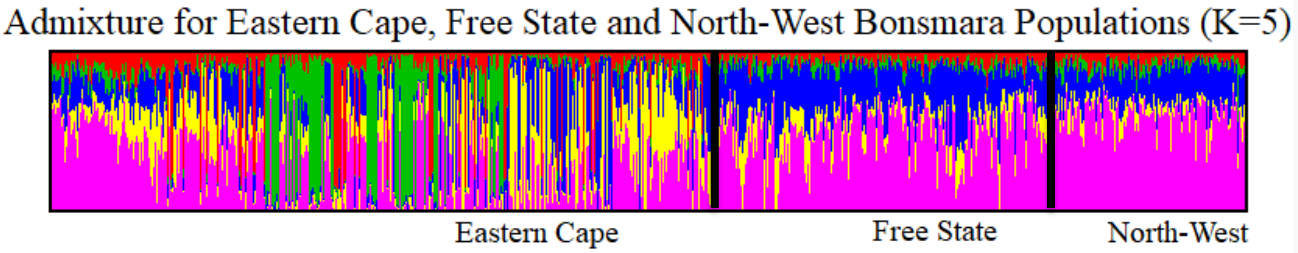


Figure 3.5 Admixture plot for Bonsmara populations from EC, FS, NW (K=5). Black lines indicate the separation of cattle by province.

The correlation strength of the environmental variables was tested using the *rstatix* package in RStudio (Kassambara, 2023). Tables 3.1, 3.2, and 3.3 exhibit the existing correlation coefficients between the annual average environmental variable recordings of each province, over the 2016-2021 periods.

The correlation coefficients for the EC environmental variables indicated a strong positive relationship between summer and winter month temperatures, and a moderate positive relationship between summer month temperature and relative humidity. A moderate negative relationship was observed between winter month temperature and relative humidity. While weak inverse relationships were observed for summer month temperature and precipitation, winter month temperature and precipitation, and relative humidity and precipitation.

Table 3.1 Correlation matrix depicting the relationships between the average annual environmental variables for the EC from 2016 to 2021.

	SMT (°C)	WMT (°C)	Rhum (%)	Precipitation (mm)
SMT (°C)	**1.000**	0.503	0.385	-0.103
WMT (°C)	0.503	**1.000**	-0.286	-0.023
Rhum (%)	0.385	-0.286	**1.000**	-0.143
Precipitation (mm)	-0.103	-0.023	-0.143	**1.000**

The correlation coefficients for the FS environmental variables (Table 3.2) indicate strong positive correlations observed between summer month and winter month temperatures, and relative humidity and precipitation.

Table 3.2 Correlation matrix depicting the relationships between the average annual environmental variables for the FS from 2016 to 2021.

	SMT (°C)	WMT (°C)	Rhum (%)	Precipitation (mm)
SMT (°C)	**1.000**	0.857	-0.344	-0.210
WMT (°C)	0.857	**1.000**	-0.131	-0.153
Rhum (%)	-0.344	-0.131	**1.000**	0.557
Precipitation (mm)	-0.210	-0.153	0.557	**1.000**

Table 3.3 indicates the correlation coefficients for the NW environmental variables. The coefficients indicate that a strong positive relationship was observed between the summer and winter month temperatures, similar to other provinces.

Table 3.3 Correlation matrix depicting the relationships between the average annual environmental variables for the NW from 2016 to 2021.

	SMT (°C)	WMT (°C)	Rhum (%)	Precipitation (mm)
SMT (°C)	**1.000**	0.858	-0.291	-0.077
WMT (°C)	0.858	**1.000**	-0.030	-0.013
Rhum (%)	-0.291	-0.030	**1.000**	0.025
Precipitation (mm)	-0.077	-0.013	0.025	**1.000**

3.2 Output results

The LEA v.3.8.0 results presented candidate loci for each respective analysis run. The number of candidate loci produced for each environmental variable and each province are summarised in Table 3.1. Notably, the LEA v.3.8.0 analysis identified very similar, if not the same (as seen in the Eastern Cape and Free State provinces), candidate loci for all the environmental variables within province. The number of detected candidate loci were similar for the all the variables of the respective provinces and when reviewing the individual loci in each list, it was observed that LEA v.3.8.0 detected the same candidate loci for the environmental variables within the provinces. Except for five loci between EC variables; four loci between FS variables; and nine loci between NW variables. Although the detected candidate loci were similar between variables within province, they differed from those detected from the environmental variables from other provinces.

Table 3.4 indicates K=1-3; these K values are based on the number of clusters observed for each province in the PCA plots (Figures 3.1 and 3.2). In Figure 3.1, the EC province has three clusters, the FS province has one cluster, and the NW province also has one cluster. Collectively, K=5, but given that LEA v.3.8.0 analysis was run for each individual province, individual K values based on their PCA genetic clustering was appropriate. The R. script used for the LEA v.3.8.0 analyses is found in

Addendum B and the visualised Manhattan plots and histograms of the candidate loci p-values are found in Addendum C.

Table 3.4 List of the number of candidate loci detected with LEA v.3.8.0 analysis.

K	Summer month temperature	Winter month temperature	Relative humidity	Precipitation
		Eastern Cape		
K = 3	Candidates: 1272	Candidates: 1267	Candidates: 1267	Candidates: 1267
		Free State		
K = 1	Candidates: 1231	Candidates: 1231	Candidates: 1227	Candidates: 1227
		North-West		
K = 1	Candidates 2177	Candidates: 2179	Candidates: 2170	Candidates: 2177

The SNP names for the candidate loci were determined by cross-referencing the candidate loci list with the SNP map file. Afterward, the F_{ST} values for all the candidate SNPs were calculated using Plink v1.90b6.18 64-bit (Purcell *et al.*, 2007) and the 20 SNPs with the highest F_{ST} values were chosen for gene annotation. Given that the majority of the candidate loci identified for the environmental variables within the provinces were very similar, the same top 20 SNPs, based on F_{ST} values, were identified within the provinces. These SNPs are listed in Table 3.5.

Table 3.5 List of SNP variants chosen for gene annotation based on F_{ST} value.

Province	SNP name	F_{ST} value
Eastern Cape	Hapmap60763-rs29014471	0.220403
	ARS-BFGL-NGS-15262	0.159269
	UA-IFASA-6154	0.136891
	Hapmap30781-BTA-129235	0.123797
	BTA-25512-no-rs	0.12186
	Hapmap52624-rs29010906	0.121411
	BovineHD0500016099	0.120668
	ARS-BFGL-NGS-95580	0.118472
	Hapmap47719-BTA-74313	0.109666
	BTA-54868-no-rs	0.107576
	BTB-00690781	0.104844
	Hapmap46867-BTA-74870	0.103043
	ARS-BFGL-NGS-42574	0.102842
	BovineHD0200037035	0.101988
	BTA-101385-no-rs	0.098452
	BTB-01216169	0.0956736
	Hapmap42094-BTA-119925	0.0953916
	ARS-BFGL-NGS-61477	0.095326
	UA-IFASA-4213	0.094439
Free State	ARS-BFGL-BAC-2790	0.237314
	ARS-BFGL-NGS-113925	0.169203
	ARS-BFGL-NGS-118306	0.157003
	BTB-00358369	0.132608
	BovineHD1800012264	0.131119
	BTA-25512-no-rs	0.12186
	ARS-BFGL-NGS-11435	0.114971
	BTA-21281-no-rs	0.114971
	Hapmap46867-BTA-74870	0.103043

	Hapmap35516-SCAFFOLD70087_15055	0.0986397
	Hapmap41178-BTA-120553	0.0939485
	Hapmap40537-BTA-43945	0.0936463
	ARS-BFGL-NGS-58613	0.0932169
	Hapmap50985-BTA-49760	0.0913317
	Hapmap41389-BTA-95970	0.089002
	ARS-BFGL-NGS-116503	0.0882002
	ARS-BFGL-NGS-99210	0.0850691
	ARS-BFGL-NGS-73895	0.084101
	BTB-00049653	0.0837741
	ARS-BFGL-NGS-113915	0.0835864
North-West	ARS-BFGL-NGS-118306	0.157003
	BovineHD0100037733	0.142343
	Hapmap30781-BTA-129235	0.123797
	ARS-BFGL-NGS-117841	0.119023
	ARS-BFGL-NGS-11435	0.114971
	Hapmap43154-BTA-11252	0.112013
	BTA-110709-no-rs	0.11135
	BovineHD1500006335	0.107096
	BTB-00690781	0.104844
	Hapmap46867-BTA-74870	0.103043
	BovineHD0200037035	0.101988
	BTB-01548042	0.100665
	ARS-BFGL-NGS-20165	0.0965003
	ARS-BFGL-NGS-27414	0.0962613
	BovineHD2200002791	0.0952757
	BTA-42967-no-rs	0.0943384
	ARS-BFGL-NGS-67912	0.0935373
	BTB-00049981	0.0922843
	BTB-01349174	0.0919727
	ARS-BFGL-NGS-107749	0.0904915

3.3 Gene annotation results

The IDs of the 60 SNPs chosen for annotation were determined from the European Variant Archive (EVA, http://www.ebi.ac.uk/eva) using the chromosome number and position of the SNPs. Table 3.6 summarises the preliminary SNP information for the respective provinces. The SNPs that were common between the provinces are in bold. One SNP was common among all three provinces; Hapmap46867-BTA-74870, found on BTA5. The Eastern Cape and the Free State cows had an additional SNP in common; BTA-25512-no-rs, found on BTA2. The Eastern Cape and the North-West cows had one additional SNP in common; BTB-00690781, found on BTA18. The Free State and the North-West had two additional SNPs in common; ARS-BFGL-NGS-118306, found on BTA1, and ARS-BFGL-NGS-11435, found on BTA10.

Table 3.6 Summary of SNP IDs and relevant genetic information (EVA, 2023).

SNP name	ID	Chromosome	Position
Eastern Cape			
Hapmap60763-rs29014471	rs29014471	4	57984488
ARS-BFGL-NGS-15262	rs111013945	23	31312887
UA-IFASA-6154	rs41590851	8	104158107
Hapmap30781-BTA-129235	**rs41742728**	**14**	**53174026**
BTA-25512-no-rs	**rs41608694**	**2**	**26140855**
Hapmap52624-rs29010906	rs29010906	2	6047203
BovineHD0500016099	rs135460220	5	56757198
ARS-BFGL-NGS-95580	rs110720564	21	66951310
Hapmap47719-BTA-74313	rs41651515	10	70736766
BTA-54868-no-rs	rs41637678	22	56085534
BTB-00690781	**rs41859545**	**18**	**2047277**
Hapmap46867-BTA-74870	**rs41654519**	**5**	**105275043**
ARS-BFGL-NGS-42574	rs42948402	11	6322937
BovineHD0200037035	**rs110480115**	**2**	**127598142**
BTA-101385-no-rs	rs41615248	6	100328689
BTB-01216169	rs42373071	20	52674980
Hapmap42094-BTA-119925	rs41619869	4	82923306
ARS-BFGL-NGS-61477	rs109922560	11	78001260
UA-IFASA-4213	rs29017131	11	21581503
Hapmap51737-BTA-50812	rs41641772	20	53757088
Free State			
ARS-BFGL-BAC-2790	rs43305347	2	41672298
ARS-BFGL-NGS-113925	rs110991778	2	49943710
ARS-BFGL-NGS-118306	**rs110230072**	**1**	**69244252**
BTB-00358369	rs43563141	8	75347512
BovineHD1800012264	rs41822969	18	41563324
BTA-25512-no-rs	**rs41608694**	**2**	**26140855**
ARS-BFGL-NGS-11435	**rs110134942**	**10**	**77234231**
BTA-21281-no-rs	rs41572236	10	95524114
Hapmap46867-BTA-74870	**rs41654519**	**5**	**105275043**
Hapmap35516-SCAFFOLD70087_15055	rs29018805	18	41533147
Hapmap41178-BTA-120553	rs41573759	2	52612052
Hapmap40537-BTA-43945	rs41582522	18	58666276
ARS-BFGL-NGS-58613	rs42254761	4	78609385
Hapmap50985-BTA-49760	rs41637019	2	129882579
Hapmap41389-BTA-95970	rs41615129	1	99642111
ARS-BFGL-NGS-116503	rs110915150	26	47230247
ARS-BFGL-NGS-99210	rs110578763	17	52740918
ARS-BFGL-NGS-73895	rs109511653	26	6948695
BTB-00049653	rs43263479	1	107910690
ARS-BFGL-NGS-113915	rs109316550	17	32911852
North-West			
ARS-BFGL-NGS-118306	**rs110230072**	**1**	**69244252**
BovineHD0100037733	rs137454952	1	132498578
Hapmap30781-BTA-129235	**rs41742728**	**14**	**53174026**
ARS-BFGL-NGS-117841	rs110442188	12	14724296
ARS-BFGL-NGS-11435	**rs110134942**	**10**	**77234231**
Hapmap43154-BTA-11252	rs29017015	4	62742887
BTA-110709-no-rs	rs41574618	3	6546124

BovineHD1500006335	rs41749003	15	24233148
BTB-00690781	**rs41859545**	**18**	**2047277**
Hapmap46867-BTA-74870	**rs41654519**	**5**	**105275043**
BovineHD0200037035	**rs110480115**	**2**	**127598142**
BTB-01548042	rs42668898	1	22720015
ARS-BFGL-NGS-20165	rs41903329	19	23561900
ARS-BFGL-NGS-27414	rs110815204	19	61205726
BovineHD2200002791	rs110088513	22	9354589
BTA-42967-no-rs	rs41581157	18	42407639
ARS-BFGL-NGS-67912	rs43713533	15	28882493
BTB-00049981	rs43266806	1	112675358
BTB-01349174	rs42468701	2	135378292
ARS-BFGL-NGS-107749	rs109830113	18	33986486

Once the IDs of the SNP variants had been identified through the European Variant Archive, they were applied to the Ensembl Variant Predictor (McLaren *et al.*, 2016) to obtain the Ensembl IDs of their respective associated genes. Table 3.7 summarises the associated gene information obtained from the Ensembl Variant Predictor (McLaren *et al.*, 2016) for the SNP variants identified from the Eastern Cape cows. 20 SNP variants were chosen for gene annotation from the Eastern Cape cows, based on their Fst values. Of those 20 chosen SNPs, 11 were identified by the Ensembl Variant Predictor (McLaren *et al.*, 2016) to be associated with genes. One of those 11 SNPs was associated with an unclassified gene. BTA-54868-no-rs was associated with gene ENSBTAG00000054247 and ENSBTAG00000035827, but gene ENSBTAG00000054247 was not recognised by the Panther Classification System database (Mi et al., 2019).

Table 3.7 Summary of SNP and gene information obtained from Ensembl Variant Predictor for the Eastern Cape (McLaren *et al.*, 2016).

SNP ID	ID	Location	Symbol	Gene ID
Hapmap52624-rs29010906	rs29010906	2:6109495-6109495	*C2H2orf88*	ENSBTAG00000026994
Hapmap52624-rs29010906	rs29010906	2:6109495-6109495	*HIBCH*	ENSBTAG00000007787
BTA-25512-no-rs	rs41608694	2:26092212-26092212	*MYO3B*	ENSBTAG00000003626
Hapmap60763-rs29014471	rs29014471	4:57623147-57623147	*IMMP2L*	ENSBTAG00000004398
UA-IFASA-6154	rs41590851	8:102459234-102459234	*FKBP15*	ENSBTAG00000005116
ARS-BFGL-NGS-42574	rs42948402	11:6290438-6290438	*RFX8*	ENSBTAG00000033998
UA-IFASA-4213	rs29017131	11:21555819-21555819	*MAP4K3*	ENSBTAG00000016442
Hapmap30781-BTA-129235	rs41742728	14:51026484-51026484	*CSMD3*	ENSBTAG00000038281
BTB-00690781	rs41859545	18:2015648-2015648	*GLG1*	ENSBTAG00000002303
Hapmap51737-BTA-50812	rs41641772	20:53680094-53680094	*CDH18*	ENSBTAG00000037844

ARS-BFGL-NGS-95580	rs110720564	21:65301050-65301050	*WDR25*	ENSBTAG00000007233
BTA-54868-no-rs	rs41637678	22:55445815-55445815	-	ENSBTAG00000054247
BTA-54868-no-rs	rs41637678	22:55445815-55445815	*ATG7*	ENSBTAG00000035827

Table 3.8 lists the Ensembl gene IDs for the SNP variants identified from the Free State cows. Not all the SNP variants were found to be associated with genes through the Ensembl Variant Predictor (McLaren *et al.*, 2016). Of the ten Free State SNP variants that were linked to genes, two were linked to unclassified genes. ARS-BFGL-NGS-113915 was associated with gene ENSBTAG00000055005, which was not recognised by Panther Classification System database (Mi *et al.*, 2019). Hapmap40537-BTA-43945 was associated with gene ENSBTAG00000053131, which was recognised by Panther Ontology as a C2H2 zinc finger transcript ion ((Mi *et al.*, 2019).

Table 3.8 Summary of SNP and gene information obtained from Ensembl Variant Predictor for the Free State (McLaren *et al.*, 2016).

SNP ID	ID	Location	Symbol	Gene ID
BTB-00049653	rs43263479	1:107060820-107060820	*SMC4*	ENSBTAG00000005862
BTA-25512-no-rs	rs41608694	2:26092212-26092212	*MYO3B*	ENSBTAG00000003626
Hapmap41178-BTA-120553	rs41573759	2:52511673-52511673	*GTDC1*	ENSBTAG00000001132
BTB-00358369	rs43563141	8:74054145-74054145	*ADRA1A*	ENSBTAG00000031632
ARS-BFGL-NGS-11435	rs110134942	10:76922630-76922630	*SPTB*	ENSBTAG00000004732
ARS-BFGL-NGS-113915	rs109316550	17:32550404-32550404	-	ENSBTAG00000055005
ARS-BFGL-NGS-99210	rs110578763	17:50485602-50485602	*TMEM132B*	ENSBTAG00000006626
Hapmap40537-BTA-43945	rs41582522	18:58328337-58328337	-	ENSBTAG00000053131
ARS-BFGL-NGS-73895	rs109511653	26:6941707-6941707	*PRKG1*	ENSBTAG00000018404
ARS-BFGL-NGS-116503	rs110915150	26:46882777-46882777	*DOCK1*	ENSBTAG00000031890

Table 3.9 presents the Ensembl gene IDs for the SNP variants identified from the North-West cows. Of the initial 20 SNP variants chosen from the F_{st} values, only ten were found to be associated with genes by the Ensembl Variant Predictor (McLaren *et al.*, 2016). As was the case with the SNPs for the Free State cows, two of the ten SNPs from the North-West cows were associated with unclassified genes. BTB-01548042 was associated with gene ENSBTAG00000048578, and this gene was not recognized by Panther Classification System database (Mi *et al.*, 2019). ARS-BFGL-NGS-67912 was associated

with gene ENSBTAG00000051802, and whilst this gene was recognised by Panther Classification System database (Mi *et al.*, 2019), it was an unclassified gene and no information was available for it.

Table 3.9 Summary of SNP and gene information obtained from Ensembl Variant Predictor for the North-West (McLaren *et al.*, 2016).

SNP ID	ID	Location	Symbol	Gene ID
ARS-BFGL-NGS-11435	rs110134942	10:76922630-76922630	*SPTB*	ENSBTAG00000004732
ARS-BFGL-NGS-117841	rs110442188	12:14680247-14680247	*TSC22D1*	ENSBTAG00000047739
ARS-BFGL-NGS-27414	rs110815204	19:60628536-60628536	*KCNJ16*	ENSBTAG00000024903
Hapmap30781-BTA-129235	rs41742728	14:51026484-51026484	*CSMD3*	ENSBTAG00000038281
BovineHD1500006335	rs41749003	15:23899610-23899610	*TTC12*	ENSBTAG00000010008
BTB-00690781	rs41859545	18:2015648-2015648	*GLG1*	ENSBTAG00000002303
ARS-BFGL-NGS-20165	rs41903329	19:22981474-22981474	*RTN4RL1*	ENSBTAG00000012302
BTB-01548042	rs42668898	1:23214124-23214124	-	ENSBTAG00000048578
BTB-00049981	rs43266806	1:111775805-111775805	*GMPS*	ENSBTAG00000013013
ARS-BFGL-NGS-67912	rs43713533	15:28453282-28453282	-	ENSBTAG00000051802

The Ensembl gene IDs that were obtained from the Ensembl Variant Predictor (McLaren *et al.*, 2016) were loaded into Panther Classification System database (Mi *et al.*, 2019) to gain more information about all the respective genes. There were ten of the 20 Eastern Cape SNPs chosen for gene annotation that were found by Panther Classification System database to be associated with genes. The basic identification information for these genes is summarised in Table 3.10. The myosin IIIB, *MYO3B*; Golgi apparatus protein 1, *GLG1*; and coiled-coil domain-containing protein 69, *CCDC69* genes were not placed in any Panther protein class.

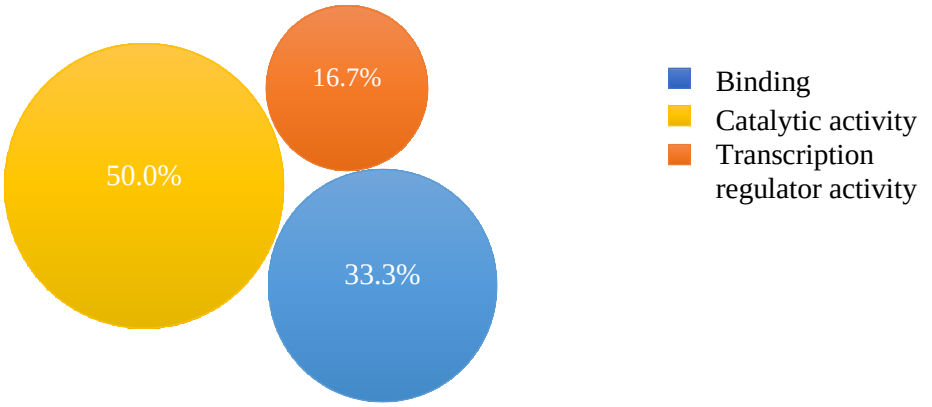
Table 3.10 Summary of the associated gene information for the Eastern Cape SNP variants (Mi *et al.*, 2019; Pantherdb, 2023b, 2023c, 2023d, 2023e, 2023g, 2023h, 2023k, 2023l, 2023o, 2023p, 2023r, 2023z).

Gene ID	Mapped IDs	Gene Name Gene Symbol	PANTHER Family/Subfamily	PANTHER Protein class
BOVIN|ENSBTAG00000026994|UniProtKB=Q3SZY8	ENSBTAG00000026994	Small membrane A-kinase anchor protein *C2H2orf88*	Small membrane A-kinase anchor protein	Scaffold/adaptor protein
BOVIN|ENSBTAG00000007787|UniProtKB=A0A3Q1LUU2	ENSBTAG00000007787	3-hydroxyisobutyryl-CoA hydrolase, mitochondrial	3-hydroxyisobutyryl-CoA	Hydrolase

		HIBCH	hydrolase, mitochondrial	
BOVIN|ENSBTAG00000003626|UniProtKB=A0A3Q1M4N6	ENSBTAG00000003626	Myosin IIIB *MYO3B*	Myosin IIIB	-
BOVIN|ENSBTAG00000004398|UniProtKB=Q2KI92	ENSBTAG00000004398	Mitochondrial inner membrane protease subunit 2 *IMMP2L*	Mitochondrial inner membrane protease subunit 2	Protease
BOVIN|ENSBTAG00000005116|UniProtKB=F1MNV8	ENSBTAG00000005116	Peptidylprolyl isomerase *FKBP15*	FK506-binding protein 15	-
BOVIN|ENSBTAG00000033998|UniProtKB=E1BAR6	ENSBTAG00000033998	RFX family member 8, lacking RFX DNA binding domain *RFX8*	DNA-binding protein RFX8	Winged helix/forkhead transcription factor
BOVIN|ENSBTAG00000016442|UniProtKB=E1BJH5	ENSBTAG00000016442	Mitogen-activated protein kinase kinase kinase kinase 3 *MAP4K3*	Mitogen-activated protein kinase kinase kinase kinase 3	Non-receptor serine/threonine protein kinase
BOVIN|ENSBTAG00000038281|UniProtKB=A0A3Q1LWF2	ENSBTAG00000038281	CUB and Sushi multiple domains 3 *CSMD3*	CUB and Sushi domain containing protein 3	-
BOVIN|ENSBTAG00000002303|UniProtKB=A0A3Q1LL67	ENSBTAG00000002303	Golgi apparatus protein 1 *GLG1*	Golgi apparatus protein 1	-
BOVIN|ENSBTAG00000037844|UniProtKB=Q08DJ5	ENSBTAG00000037844	Cadherin-18 *CDH18*	Cadherin-18	Cadherin
BOVIN|ENSBTAG00000007233|UniProtKB=A0A3Q1NBM5	ENSBTAG00000007233	WD repeat domain 25 *WDR25*	WD repeat containing protein 25	-
BOVIN|ENSBTAG00000035827|UniProtKB=A0A3Q1LXA9	ENSBTAG00000035827	Ubiquitin-like modifier-activating enzyme ATG7 *ATG7*	Ubiquitin-like modifier-activating enzyme ATG7	Ubiquitin-protein ligase

Figures 3.6(a) and 3.6(b) depict the distributions of the types of molecular functions and biological processes for the classified genes identified from the EC cow LEA v.3.8.0 analyses. The molecular functions included binding (33.3%), catalytic activity (50.0%), and transcription regulator activity (16.7%). The biological processes included biological adhesion (5.3%), biological regulation (10.5%), cellular processes (31.6%), developmental processes (10.5%), localization (5.3%), metabolic processes (21.1%), multicellular organismal processes (10.5%), and response to stimulus (5.3%). The unclassified genes to which no PANTHER category was assigned were not included in the figures.

(a)



(b)

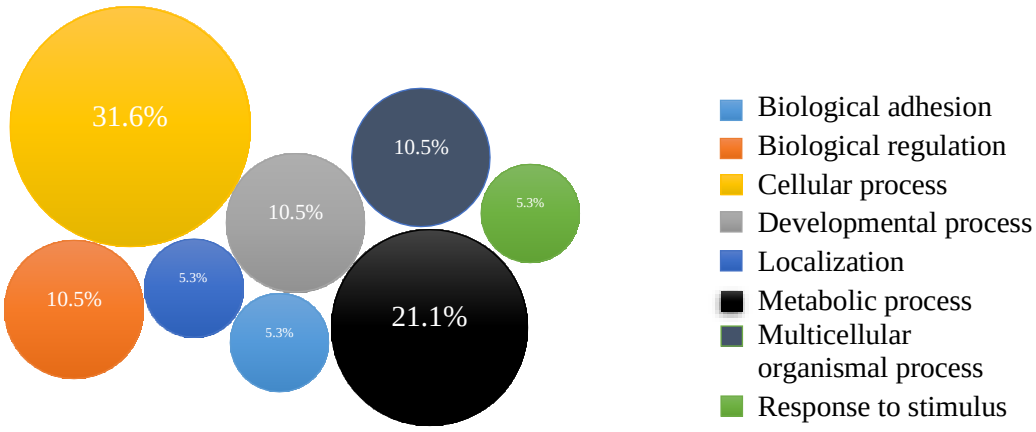


Figure 3.6 The proportions of the molecular functions (a) and biological processes (b) for the genes identified from the EC cow analyses (Mi *et al.*, 2019).

There were 20 SNPs chosen for gene annotation from the Free State. Nine of the 20 SNPs were found by Panther Classification System database to be associated with genes (Mi *et al.*, 2019). Table 3.11 summarises the basic identification information for the nine genes. Three of the nine genes did not fall into a Panther protein class. Notably, the Free State and Eastern Cape cows share the Myosin IIIB (*MYOB3*) gene.

Table 3.11 Summary of the associated gene information for the Free State SNP variants (Mi *et al.*, 2019; Pantherdb, 2023a, 2023f, 2023j, 2023n, 2023p, 2023q, 2023u, 2023v, 2023w).

Gene ID	Mapped IDs	Gene Name Gene Symbol	PANTHER Family/Subfamily	PANTHER Protein class
BOVIN|ENSBTAG00000005862|UniProtKB=E1BMZ9	ENSBTAG00000005862	Structural maintenance of chromosomes protein *SMC4*	Structural maintenance of chromosomes protein 4	-
BOVIN|ENSBTAG00000003626|UniProtKB=A0A3Q1M4N6	ENSBTAG00000003626	Myosin IIIB *MYO3B*	Myosin IIIB	-

BOVIN|ENSBTAG00000001132|UniProtKB=Q08DA7	ENSBTAG00000001132	Glycosyltransferase-like domain-containing protein 1 *GTDC1*	Glycosyltransferase-like domain-containing protein 1	Glycosyltransferase
BOVIN|ENSBTAG00000031632|UniProtKB=A0A3Q1MAT2	ENSBTAG00000031632	Alpha-1A adrenergic receptor *ADRA1A*	Alpha-1A adrenergic receptor	G-protein coupled receptor
BOVIN|ENSBTAG00000004732|UniProtKB=F1MKE9	ENSBTAG00000004732	Spectrin beta chain *SPTB*	Spectrin beta chain, erythrocytic	Actin or actin-binding cytoskeletal protein
BOVIN|ENSBTAG00000006626|UniProtKB=E1BKR8	ENSBTAG00000006626	Transmembrane protein 132B *TMEM132B*	Transmembrane protein 132B	-
BOVIN|ENSBTAG00000053131|UniProtKB=A0A3Q1LXZ5	ENSBTAG00000053131	KRAB domain-containing protein -	KRAB domain-containing protein ZNF747-related	C2H2 zinc finger transcription factor
BOVIN|ENSBTAG00000018404|UniProtKB=P00516	ENSBTAG00000018404	cGMP-dependent protein kinase 1 *PRKG1*	-	Non-receptor serine/threonine protein kinase
BOVIN|ENSBTAG00000031890|UniProtKB=F1MIX6	ENSBTAG00000031890	Dedicator of cytokinesis 1 *DOCK1*	Dedicator of cytokinesis protein 1	Guanyl-nucleotide exchange factor

The molecular function and biological process types and distributions for the genes identified from the FS cow LEA v.3.8.0 analyses are presented in Figures 3.7(a) and 3.7(b). The molecular functions were grouped into four types: binding (40.0%), catalytic activity (20.0%), molecular function regulation (20.0%), and molecular transducer activity (20.0%). The biological processes were grouped into six types. 33.3% of the classified genes were involved in cellular processes, and 12.5% were involved in biological regulation, localization, locomotion, response to stimulus, and signalling, respectively. The unclassified genes to which no PANTHER category was assigned were not included in the figures.

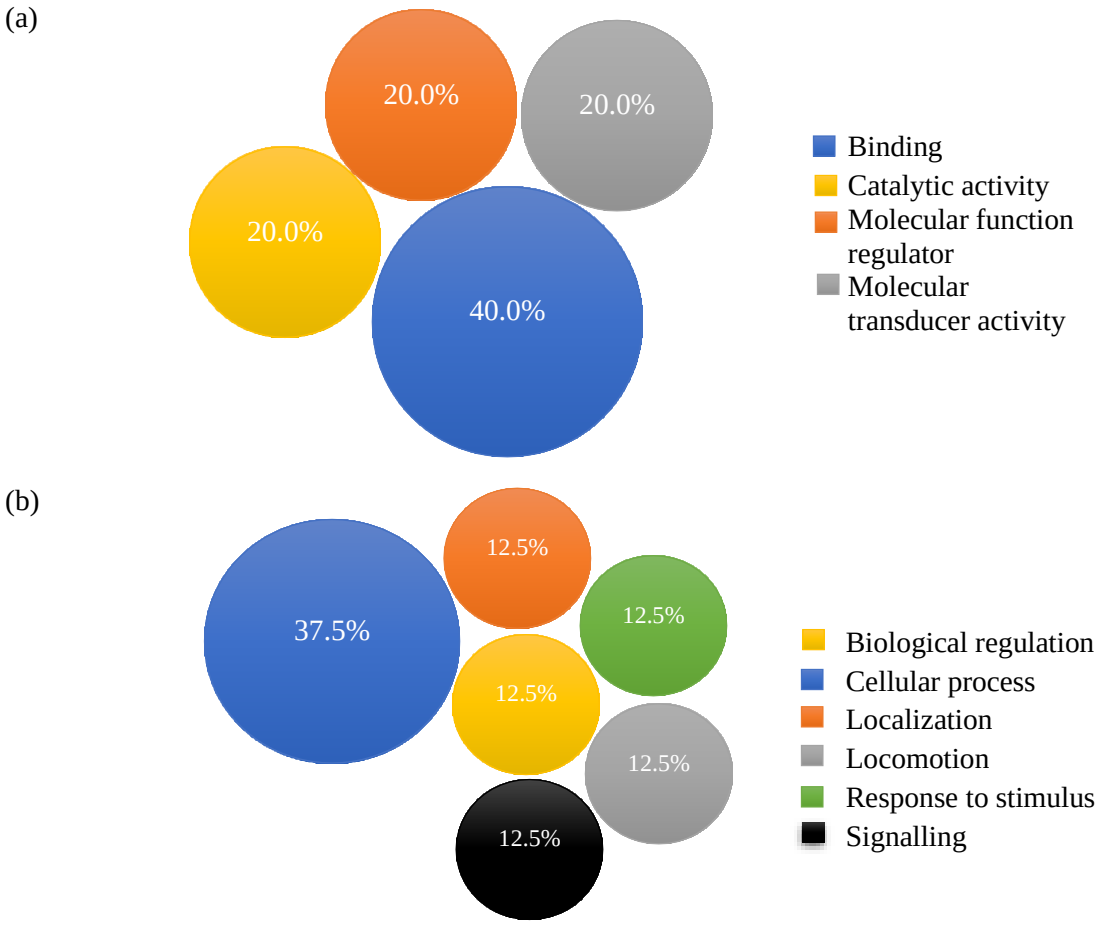


Figure 3.7 The proportions of the molecular functions (a) and biological processes (b) for the genes identified from the FS cow analyses (Mi *et al.*, 2019).

Of the original 20 North-West SNPs chosen for gene annotation, eight were found by Panther Classification System database to be associated with genes (Mi *et al.*, 2019). The basic identification information for these genes is summarised in Table 3.12. Five of the eight genes did not form part of a Panther protein class. The *SPTB*, the spectrin beta chain, erythrocytic gene, was detected in both the Free State and North-West cows. The Golgi apparatus protein 1 (*GLG1*) and the CUB and Sushi domain-containing protein 3 (*CSMD3*) genes were detected in both the Eastern Cape and North-West cows.

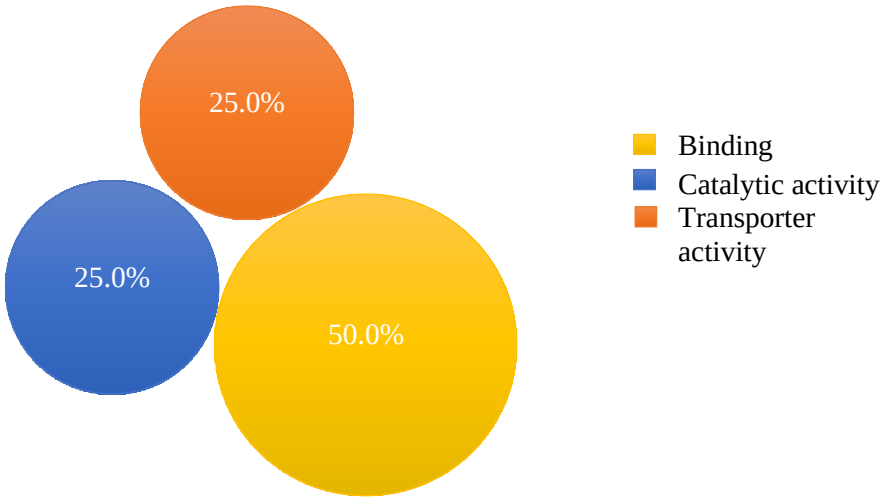
Table 3.12 Summary of the associated gene information for the North-West SNP variants (Mi *et al.*, 2019; Pantherdb, 2023e, 2023h, 2023i, 2023m, 2023s, 2023v, 2023x, 2023y).

Gene ID	Mapped IDs	Gene Name Gene Symbol	PANTHER Family/Subfamily	PANTHER Protein class
BOVIN|ENSBTAG00000004732|UniProtKB=F1MKE9	ENSBTAG00000004732	Spectrin beta chain *SPTB*	Spectrin beta chain, erythrocytic	Actin or actin-binding cytoskeletal protein

BOVIN|ENSBTAG00000010008|UniProtKB=E1BEU1	ENSBTAG00000047739	TSC22 domain family protein 1 *TSC22D1*	TSC22 domain family protein 1	-
BOVIN|ENSBTAG00000024903|UniProtKB=Q0VD28	ENSBTAG00000024903	Potassium inwardly rectifying channel subfamily J member 16 *KCNJ16*	Inward rectifier potassium channel 16	Ion channel
BOVIN|ENSBTAG00000038281|UniProtKB=A0A3Q1LWF2	ENSBTAG00000038281	CUB and Sushi multiple domains 3 *CSMD3*	CUB and Sushi domain-containing protein 3	-
BOVIN|ENSBTAG00000047739|UniProtKB=Q3MHL6	ENSBTAG00000010008	Tetratricopeptide repeat domain 12 *TTC12*	Tetratricopeptide repeat protein 12	-
BOVIN|ENSBTAG00000002303|UniProtKB=A0A3Q1LL67	ENSBTAG00000002303	Golgi apparatus protein 1 *GLG1*	Golgi apparatus protein 1	-
BOVIN|ENSBTAG00000012302|UniProtKB=E1BEI7	ENSBTAG00000012302	Reticulon 4 receptor like 1 *RTN4RL1*	Reticulon-4-receptor-like-1	-
BOVIN|ENSBTAG00000013013|UniProtKB=A0A3Q1LUN6	ENSBTAG00000013013	Glutamine amidotransferase *GMPS*	GMP synthase [glutamine-hydrolysing]	Ligase

The types and distributions of molecular function and biological processes relevant to the classified genes identified from the NW cow LEA v.3.8.0 analyses are illustrated in the Figures 3.8(a) and 3.8(b). There were three different molecular functions that the genes were involved in: binding (50.0%), catalytic activity (25.0%), and transporter activity (25.0%).

(a)



Eight different biological processes were illustrated in Figure 3.8(b). Most of the genes were again involved in cellular processes (33.3%). Biological regulation, developmental processes, and multicellular organismal processes, all had 13.3% of the identified classified genes involved with them. Localization, locomotion, metabolic processes, and response to stimuli, all had 6.7% of the identified classified genes involved with them. The unclassified genes to which no PANTHER category was assigned were not included in the figures.

(b)

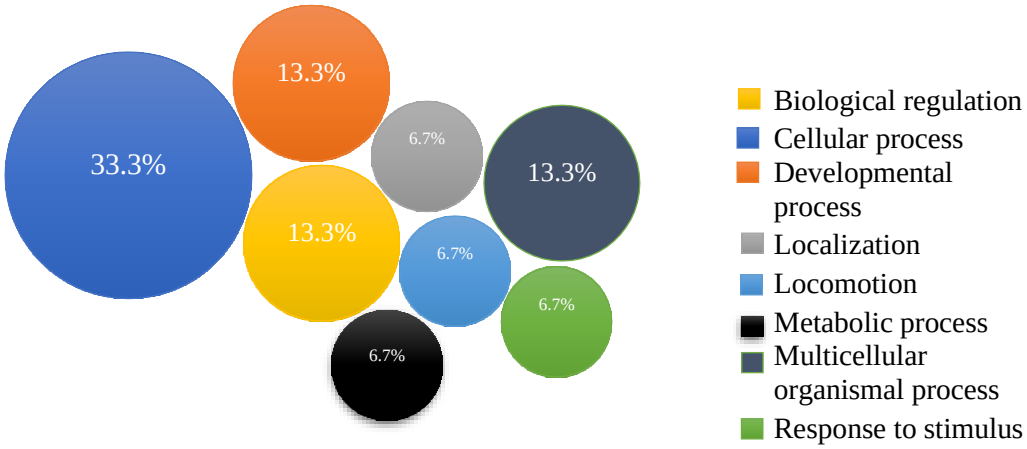


Figure 3.8 The proportions of the molecular functions (a) and biological processes (b) for the genes identified from the NW cow analyses (Mi *et al.*, 2019).

3.4 GWAS

The GWAS results were visualised through Manhattan plots using various R packages; *tidyverse*, *ggtext* and *ggplot* (Purcell *et al.*, 2007; Wickham, 2016; Wickham *et al.*, 2019; Wilke & Wiernik, 2022). The significance level of the Manhattan plots was set to 10^{-8}, based on suggestion that this is considered the significance level of a convincing association (Pirinen, 2023). The results from the GWAS for birth weight and the set of 25 272 common SNPs, which were used for the landscape genomics LEA v.3.8.0 analyses, are visualised in Figure 3.9. No significant SNP associations were identified for this trait.

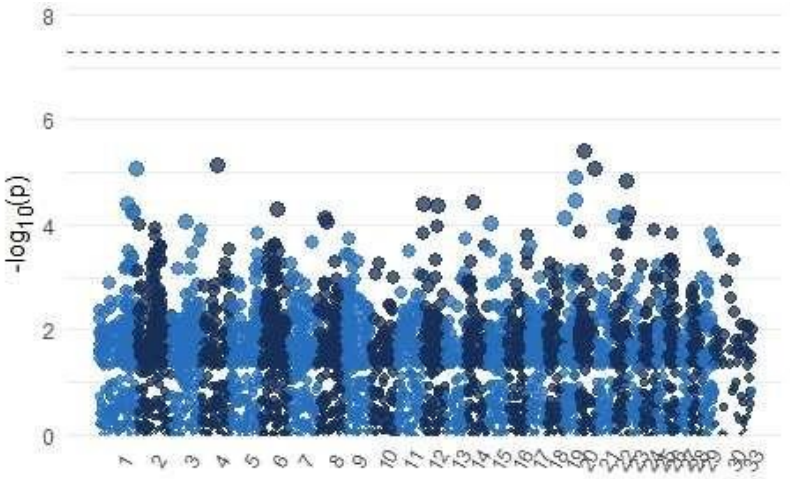


Figure 3.9 Manhattan plot for the significant SNPs associated with BW, derived from the 25 272 common SNPs also used in LEA analysis.

Figure 3.10 is the Manhattan plot that visualises the GWAS results for weaning weight and the set of 25 272 common SNPs, which were used for the landscape genomics LEA v.3.8.0 analyses. There were no significant SNP associations identified for this trait.

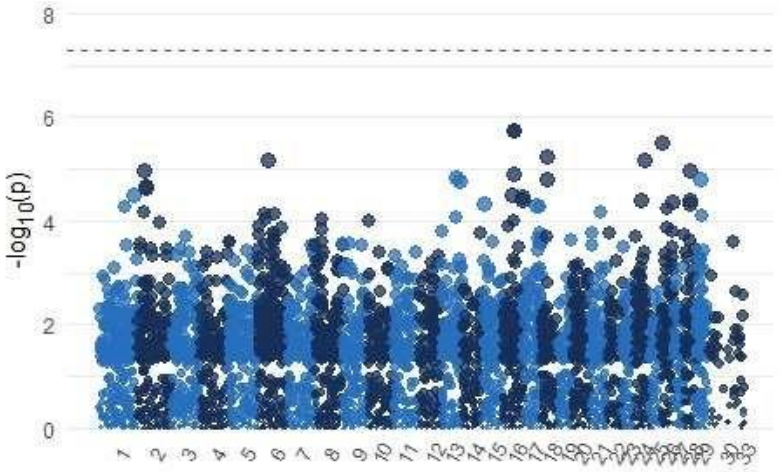


Figure 3.10 Manhattan plot for the significant SNPs associated with WW, derived from the 25 272 common SNPs also used in LEA analysis.

The results from the GWAS for yearling weight and the set of 25 272 common SNPs, which were used for the landscape genomics LEA v.3.8.0 analyses, are visualised in the Manhattan plot Figure 3.11. There were also no significant SNP associations identified for this trait.

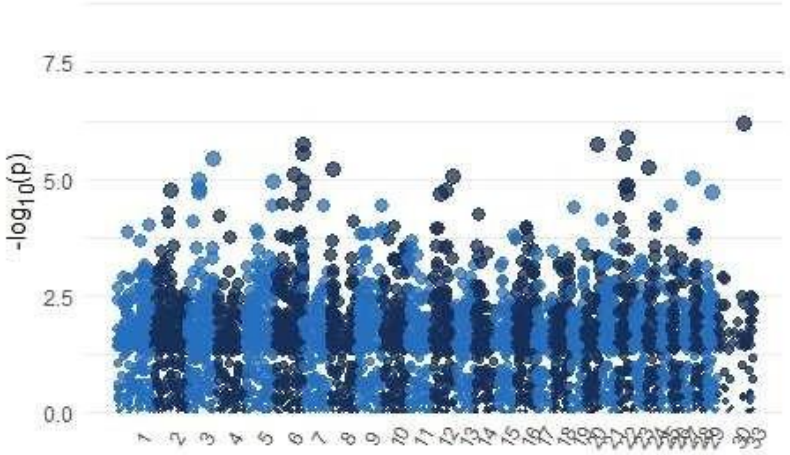


Figure 3.11 Manhattan plot for the significant SNPs associated with YW, derived from the 25 272 common SNPs also used in LEA analysis.

The GWAS results for eighteen-month weight and the set of 25 272 common SNPs, used for the landscape genomics LEA v.3.8.0 analyses, are visualised in Figure 3.12. These results indicated four significant SNP associations with the growth trait. The first three SNPs (Hapmap49016-BTA-110674, Hapmap32099-BTA-151095, Hapmap59651-rs29009956) are located on BTA6, and the fourth (ARS-BFGL-NGS-26337) is located on BTA26.

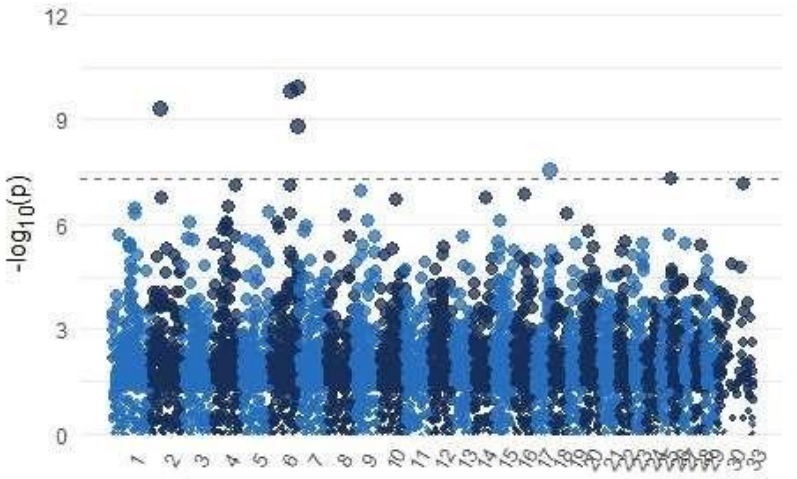


Figure 3.12 Manhattan plot for the significant SNPs associated with 18MW, derived from the 25 272 common SNPs also used in LEA analysis.

The SNPs with significant associations with 18MW observed in Figure 3.12 were compared to the significant candidate SNP loci identified by the LEA v.3.8.0 analysis. There were no similarities between the results.

Chapter 4: Discussion

4.1 Introduction

The Bonsmara breed is considered the most popular composite beef breed in South Africa with an estimated 130 000 registered stud cattle (Brand *et al.*, 2021; UP George Bonsmaras, 2023). The breed is known for its adaptability to the diverse SA climate and is also farmed in countries such as Namibia, Botswana, and South America (Strydom *et al.*, 2016; Gouws, 2020; Barbosa *et al.*, 2022). Previous studies (Visagie, 2012; Webb *et al.*, 2017) confirmed the role of the environment on the growth and performance of Bonsmara cattle. This study aimed to investigate the environmental association with growth in Bonsmara based on genotypic information. A landscape genomic approach was followed to seek insight on the role of the environment on growth performance. 18MW was the most suitable phenotype due to the available recordings.

Growth performance is associated with various economically important traits that are important to commercial farmers (Zhuang *et al.*, 2020). In addition, the associated traits (BW, WW, YW, 18MW) are important for stud farmers to consider as bulls and cows need to meet specific breed standards to be registered with breed societies (Bonsmara SA, 2019). In this study, growth performance in Bonsmara cows from different South African regions including the Eastern Cape, Free State, and North-West provinces was investigated. It should be noted that for a landscape genomics analysis, the animals included had to be born and grown in the same environment. Although, 4769 genotypes were available, after careful investigation only 765 Bonsmara cows met the criteria, having 18MW measurement and lived their whole productive life on the same farm and environment.

Summer and winter month temperatures, relative humidity, and rainfall were analysed alongside the genotypes of 765 Bonsmara cows (following quality control) using a landscape genomics approach through the LEA v.3.8.0 package (Frichot & Francois, 2015a) in RStudio (Rstudio, 2022). The candidate loci that were produced from the LEA v.3.8.0 analysis were used to identify and annotate genes that were potentially associated with growth traits. Landscape genomics is a relatively new approach in the livestock sector (Joost & Negrini, 2010). It is also a rapidly evolving field with methodologies being consistently changed and adapted as novel genomic information about species is discovered (Storfer *et al.*, 2018).

4.2 Population structure analysis

The first step in this study was to consider the population structure of the cows representing the three provinces. The PCA plots confirmed the genetic variability and being from the same breed, the cows clustered together near the centre of the axes. The two EC clusters indicated two lines and cross referencing their IDs to the breeder they belong to, found that all the animals belonged to the same breeder. Thus, the separate clusters are probably due to different breeding objectives and selection

emphasis implemented by the breeder. Similar clustering for SA Bonsmara cattle was reported by Bosman *et al.* (2017) using genotypes from the same database.

In order to estimate the most correct K value for the LEA v.3.8.0 analysis, the population structure analysis was confirmed with Admixture (Pritchard *et al.*, 2000). The EC cows indicated three genetic sub clusters, while FS and NW had one each. A population structure analysis, conducted by Makina *et al.* (2016), that allowed for three ancestral populations, predicted that the Bonsmara breed composition consisted of 41% European taurine, 42% African taurine, and 16% indicine. These findings correspond to the breeds that were used to develop the Bonsmara breed, 5/8 Afrikaner, 3/16 Hereford, and 3/16 Shorthorn (Bonsma, 1985; van Marle-Köster *et al.*, 2021). This explains the consistency in admixture proportions for the individual cows observed in the population analysis in this study. The outlier groups of Eastern Cape cattle seem to have differing proportions of admixture in comparison to the other animals, and this remains constant in all three admixture plots. This observed differentiation is typical in seed stock farming where a relatively small percentage of breeders tend to supply the majority of superior genetics. The effect of this in the Bonsmara populations is high genetic linkage observed in some herds and fragmented genetic linkage in others (Bosman *et al.*, 2017).

4.3 Climate correlations

The majority of South Africa is classified as having a subtropical climate and experiencing spring and summer precipitation (September to February), with the exception of southwestern South Africa which experiences winter precipitation (Blamey *et al.*, 2018). The results showed positive correlation coefficients between summer and winter month temperatures for all three provinces, which are similar to those reported in previous literature (Irwin & Good, 2012; Chikosi *et al.*, 2018; van der Walt & Fitchett, 2021). Warmer summers can correspond to warmer winters and cooler summers can correspond to cooler winters, especially with the extreme warm temperature trends observed in South Africa (van der Walt & Fitchett, 2021). The correlation coefficients for the relationships between relative humidity and precipitation, over the 2016 to 2021 period, for the FS and NW provinces were positive. This observed relationship supports the expected relationship between relative humidity and precipitation observed in sub-tropical climates (Park & Min, 2017; Denson *et al.*, 2021).

The FS and NW provinces fall into the subtropical class, whilst the EC province is situated in the transition zone between summer precipitation regions and winter precipitation regions (Blamey *et al.*, 2018). The EC region is thus influenced by subtropical and mid latitude weather systems (Mahlalela *et al.*, 2020). The expected weather patterns for the FS and NW are cold, dry winters and hot, wet summers (Beck *et al.*, 2018). The expected weather patterns for the EC are warm summers, cold winters, and year-round rainfall although, higher precipitation has been historically recorded for the summer months (Jury, 2013; Weldon & Reason, 2014; Beck *et al.*, 2018; Engelbrecht *et al.*, 2018; Mahlalela *et al.*, 2020). Due to the subtropical climates in the provinces that were investigated, the expected relationship

between temperature and relative humidity for all three provinces is positive, with higher temperatures corresponding to higher percentages of relative humidity (Jury, 2013; Beck *et al.*, 2018). The relationship between relative humidity and precipitation is also positive (Park & Min, 2017; Singh *et al.*, 2019; Denson *et al.*, 2021).

The observed negative correlation coefficients for the relationship between summer month temperature and relative humidity, however, were not as expected for the FS and NW provinces. This could be explained by the recent droughts (2015 to 2016 and 2018 to 2020) experienced in South Africa (Mahlalela *et al.*, 2020; Meza *et al.*, 2021). In South Africa, temperatures are reported to be increasing, however, some of the average annual temperatures in the country during the 2019 to 2021 period were cooler than in past years (Mahlalela *et al.*, 2020; South African Weather Service, 2020; South African Weather Service, 2021; South African Weather Service, 2022). Cooler temperatures are likely to have limited the amount of humid air brought in by the subtropical air masses (Kalahari high pressure cell), ultimately lowering the average annual precipitation levels, contributing to the effects of the recorded droughts (Wright *et al.*, 2019; Mahlalela *et al.*, 2020). Mahlalela *et al.* (2020) reported that the east/northeast region (which forms part of the larger eastern South Africa region along with the eastern Free State) has been statistically showing a significant decline in spring precipitation. Thus, despite increasingly warmer temperatures in comparison to previous years, mildly cooler temperatures recorded during 2019 to 2021 may have resulted in the negative correlation observed between the summer and winter month temperatures and relative humidity and ultimately, less average annual precipitation as well.

The EC correlation coefficient for summer month temperature and relative humidity was positive, as expected for the EC climate. However, the correlation coefficient for winter month temperature and relative humidity was negative, as with the FS and NW provinces. Additionally, the correlation coefficient for the relationship between relative humidity and precipitation was also not as expected for the EC, as this was negative. These observations may be due to the ENSO (El-Niño Southern Oscillation), SIOD (South Indian Ocean subtropical dipole), and Southern Annular Mode ($SAM_{(2)}$) climate modes, which are very impactful on regional precipitation in South Africa (Mahlalela *et al.*, 2020). With ENSO and SIOD strongly influencing subtropical southern African precipitation in the summer (Reason, 2001; Mahlalela *et al.*, 2020), and $SAM_{(2)}$ having strong influence over western South African precipitation during the winter (Reason & Rouault, 2005). Due to regions of the EC being situated in the transition zone between the ENSO and SIOD, and $SAM_{(2)}$ climate modes, minor oscillations in any one of their patterns could substantially impact seasonal rainfall in the EC (Mahlalela *et al.*, 2020).

Furthermore, the complexity of the weather data used for this study may also have influenced the correlation coefficients as well. The limited range of data recordings and the incompleteness of the full

weather records did not provide a clear picture of historic weather patterns in the EC, FS, and NW provinces. The gaps in the data may have skewed the correlation coefficients, giving insight into only the most recent weather events, and not the province climates as a whole.

4.4 LEA v.3.8.0 analysis, gene annotation and GWAS

Summer month temperature, winter month temperature, relative humidity, and average annual precipitation were run separately, with the 25 272 SNPs, for each of the three provinces in LEA v.3.8.0 analysis and resulted in a list of candidate loci. The environmental variables that resulted in the highest number of SNP associations were summer month temperature for EC, summer and winter month temperatures for FS, and winter month temperature for NW. This indicates that temperature may be the environmental variable mostly associated with performance compared to relative humidity and precipitation (Rashamol *et al.*, 2020). Temperature is a likely variable due to the availability and accuracy of recordings based on weather bureau data. Relative humidity and precipitation are dependent on each other, with lower relative humidity associated with lower rainfall (Park & Min, 2017; Denson *et al.*, 2021).

Due to the large number of candidate loci that were detected the F_{ST} values were calculated and the top 20 from each list were chosen for gene annotation (Gugger *et al.*, 2021), to ensure that false positives are not included. False positives have been warned against for landscape genomics studies that are reliant on genome scans (Lotterhos & Whitlock, 2015; Rellstab *et al.*, 2015; Storfer *et al.*, 2018). In literature replication is regarded as a guideline for excluding false positives as loci that are repeatedly detected are less likely to be the result of confounding population structure effects or environmental covariances (Rellstab *et al.*, 2015; Storfer *et al.*, 2018). Further, literature has indicated that loci detected by multiple analysis methods or candidate genes associated with phenotypes known to be under selection, are less likely to be false positives, compared to those without a known function (Storfer *et al.*, 2018).

The following SNP variants were not included in the latter stages of the gene annotation as they were not associated with any classified genes, namely BTA-54868-no-rs, detected in the Eastern Cape cows; ARS-BFGL-NGS-113915 and Hapmap40537-BTA-43945, detected in the Free State cows; and BTB-01548042 and ARS-BFGL-NGS-67912, detected in the North-West cows. Of the seven SNP variants that were detected in more than one province, Hapmap46867-BTA-74870, BovineHD0200037035, and ARS-BFGL-NGS-118306 were not associated with any genes (classified or unclassified). No further gene information was found for them. Hapmap46867-BTA-74870 was the only SNP variant to be detected in cows from all three provinces by the LEA v.3.8.0 analysis, however, it was not located close to any of the other identified genes or pathways.

The 25 genes identified from this study were *MYO3B*, *GLG1*, *SPTB*, *C2H2orf88*, *HIBCH*, *IMMP2L*, *RFX8*, *CDH18*, *ATG7*, *MAP4K3*, *FKBP15*, *WDR25*, *SMC4*, *TMEM132B*, *GTDC1*, *ADRA1A*, *KRAB domain-containing protein*, *PRKG1*, *DOCK1*, *TSC22D1*, *KCNJ16*, *CSMD3*, *TTC12*, *RTN4RL1*, and *GMPS*. The following 17 were all confirmed to the bovine species: *MYO3B*, *SPTB*, *HIBCH*, *IMMP2L*, *CDH18*, *ATG7*, *MAP4K3*, *TMEM132B*, *GTDC1*, *ADRA1A*, *KRAB domain-containing protein*, *PRKG1*, *DOCK1*, *TSC22D1*, *KCNJ16*, *CSMD3*, and *RTN4RL1*. Of those that have been confirmed in cattle, the following nine genes have been reported to be involved with cattle growth, performance and adaptation, and are thus of particular interest to the objective of the study: *CSMD3* (Ghoreishifar *et al.*, 2020), *CDH18* (Ahmad *et al.*, 2023), *MAP4K3* (Adjei-Fremah *et al.*, 2018), *HIBCH* (Aliloo *et al.*, 2020; Kenny *et al.*, 2022), *ATG7* (Nakanishi *et al.*, 2019; Silva *et al.*, 2022), *GTDC1* (Bolormaa *et al.*, 2011), *ADRA1A* (Hromádková *et al.*, 2020), *PRKG1* (Lonergan *et al.*, 2010; Sherman *et al.*, 2010; Taye *et al.*, 2017; Vajana *et al.*, 2018), and *KCNJ16* (Sammad *et al.*, 2022). Where *CSMD3*, *CDH18*, *MAP4K3*, *HIBCH*, and *ATG7* were detected in the ES cows; *PRKG1*, *GTDC1*, and *ADRA1A*, were detected in the FS cows; and *KCNJ16* and *CSMD3*, were detected in the NW cows.

There were eight out of the 25 identified genes, across the three provinces, whose functions have yet to be confirmed in beef cattle. The main functions of these genes have been stated but they will not be discussed in detail as the genes that have been confirmed in the bovine species were the focus of this study. The eight unconfirmed genes included *GLG1*, *RFX8*, *C2H2orf88*, *FKBP15*, *WDR25*, *SMC4*, *TTC12*, and *GMPS*. In humans the Golgi apparatus protein 1, *GLG1*, gene's function has been related to the negative regulation of transforming growth factor beta receptor signaling pathway (GeneCards, 2023a). This gene was not placed in any specific Panther molecular function category. The DNA-binding protein gene, *RFX8*, was classified in the "binding" and "transcription regulator activity" metabolic function classes. Its molecular functions include the binding and regulation of genes transcribed by RNA polymerase II in the cis-regulatory region (Mi *et al.*, 2019; Pantherdb, 2023r). The small membrane A-kinase anchor protein, *C2H2orf88*, was not characterized by any specific Panther molecular function category, however, it is classified as a scaffold/adaptor protein and these proteins are generally involved in forming multi-protein complexes that combine cAMP signaling with other pathways (Carnegie *et al.*, 2009). The peptidylprolyl isomerase binding protein 15 gene, *FKBP15*, was not placed into any specific Panther molecular function category, however it is involved in actin and protein binding, as well as endocytosis (NCBI, 2023a), while the WD repeat containing protein 25 gene, *WDR25*, has been reported to be involved in cellular processes, such as apoptosis and gene regulation (GeneCards, 2023c). *WDR25* was not placed into any specific Panther molecular function category. *SMC4*, the structural maintenance of chromosomes protein 4 gene, was not characterized by any specific Panther molecular function category. This gene codes for proteins that are essential to successful chromosome transmission during replication and segregation of the genome (Harvey *et al.*, 2002). The tetratricopeptide repeat domain 12 gene, *TTC12*, was not placed in any specific Panther molecular

function category, however it is considered an encoded protein and has been found to be involved in protein complex formation (Xu *et al.*, 2015), while glutamine amidotransferase, *GMPS*, was classified with the "catalytic activity" molecular function and enables ligase activity (Mi *et al.*, 2019; Pantherdb, 2023i). All of the above mentioned genes are associated with protein pathways, however more research is needed to determine their specific functions in relation to cattle, specifically the Bonsmara breed. Whole genome sequencing could be utilized to investigate the influence of these genes in the bovine genome.

Eastern Cape

There were two genes identified from the EC cow population that have previously been confirmed in dairy cattle but were not reported to be associated with growth or adaptation traits - *IMMP2L* and *MYO3B* (Bermingham *et al.*, 2014; Skebiel *et al.*, 2018). Neither of these genes were characterized by a specific Panther molecular function category. The mitochondrial inner membrane protease subunit 2 gene, *IMMP2L*, is classified as a protease protein and it is involved in processing proteins targeted for mitochondrial compartments and the assembly of mitochondrial respiratory chain complexes (Mi *et al.*, 2019; Pantherdb, 2023l; Skibiel *et al.*, 2018). The bovine myosin IIIB, *MYO3B* gene is involved in the regulation of immune response (Chen *et al.*, 2016). It has been identified as a QTL related to resistance to bovine tuberculosis (Bermingham *et al.*, 2014).

The following five genes, detected from the EC cows, were found to be associated with growth performance: *CSMD3*, *CDH18*, *MAP4K3*, *HIBCH*, and *ATG7*. *CSMD3* was not placed in any Panther molecular function category, whilst *CDH18* was placed into the Panther molecular function "binding" class, and *MAP4K3*, *HIBCH*, and *ATG7* were placed into the molecular function "catalytic activity" class (Mi *et al.*, 2019). These classes distinguish between the activity that the product of the gene possesses at a molecular level. The product of *CDH18* would thus participate in binding activities, whilst the products of *MAP4K3*, *HIBCH* and *ATG7* would all participate in catalytic activities (Mi & Thomas, 2009).

The CUB and Sushi multiple domains 3 gene, *CSMD3*, encodes a transmembrane protein (Mi *et al.*, 2019; Pantherdb, 2023e) and has previously been reported to be involved in bovine body size and stature (Ghoreishifar *et al.*, 2020). The detection of the gene in the EC cows could be due the targeted selection for growth in Bonsmara cattle, which led to increased expression of the gene. Additionally, the *CSMD3* detection could also be due to the gene's close location on BTA14 to another, *LOC781881*, reported by Bolormaa *et al.* (2011) to be associated with intramuscular fat (IMF) content. Targeted selection by breeders for IMF may have resulted in increased expression of *LOC781881*, and subsequently *CSMD3*, due to existing linkage disequilibrium (LD) between the genes (Bush & Moore, 2012). Further, associations have been reported between 18MW and bioregion (Webb et al., 2017).

The Cadherin-18 gene, *CDH18*, is classified as a cadherin protein. Its molecular functions include cadherin binding and calcium ion binding. Furthermore, its biological processes include multicellular organism development and cell morphogenesis (Mi *et al.*, 2019; Pantherdb, 2023d). It has been reported to be significantly associated with AFC (Ahmad *et al.*, 2023). AFC is negatively correlated with body weight at different ages, meaning selection for heavier body weight creates a favourable condition for earlier reproduction (Bourdon & Brinks, 1982; da Gama *et al.*, 2021). The detection of this gene in the EC cows could explain the targeted selection of growth traits in the Bonsmara breed.

The mitogen-activated protein kinase kinase kinase kinase 3 gene, *MAP4K3*, is classified as a non-receptor serine/threonine protein kinase is a transcription regulator (Kulkarni *et al.*, 2022) and enables the catalysis of serine and threonine reactions. Kulkarni *et al.* (2022) reported that when supplemented with probiotics, *MAP4K3* gene expression is modulated and may have a positive impact on growth performance. Probiotic supplementation is common practice in the livestock industry. The positive effects observed in cattle specifically, post probiotic supplementation include enhanced growth performance, improved host innate immunity, reduced stress, and maintenance of a consistent, continuous supply of lactic acid to the rumen microbiome (Nocek *et al.*, 2002; Adjei-Fremah *et al.*, 2018). The influence and enhanced effects of probiotics are partly the result of produced antibacterial substances that combat the pathogenic microbes found in cattle rumens (Dhama *et al.*, 2008; Adjei-Fremah *et al.*, 2018). Given that probiotics are given to beef cattle to improve production performance and lower stress levels (Carraud, 1990; Adjei-Fremah *et al.*, 2018; Kelsey & Colpoys, 2018), prolonged use may have led to the modulated gene expression of *MAP4K3* in the EC Bonsmara population.

The mitochondrial 3-hydroxyisobutyryl-CoA hydrolase, *HIBCH*, is classified as a hydrolase protein. Its molecular function involves the catalysis of a reaction whereby a thioester bond is hydrolysed (Mi *et al.*, 2019; Pantherdb, 2023k). The gene's biological processes include carboxylic acid catabolic processes, organonitrogen compound catabolic processes, branched-chain amino acid metabolic processes, and alpha-amino acid metabolic processes (Mi *et al.*, 2019; Pantherdb, 2023k). All these biological processes include the chemical reactions and pathways either resulting in the breakdown of the respective molecules or involving them (Mi *et al.*, 2019; Pantherdb, 2023k). Further, a selection signature study on Kenyan dairy cattle by Aliloo *et al.* (2020), reported a candidate gene that was found in a region overlapping the *HIBCH* gene. The *HIBCH* gene was found to be near the bovine myostatin gene, which plays a critical role in bovine muscle development (Sharma *et al.*, 1999). Additionally, a study by Kenny *et al.* (2022) reported *HIBCH* to be found in a QTL region defined for carcass fat. The identification of the gene from the EC cows may be attributed to the targeted selection of the *MSTN* gene to prevent the incidence of double muscling in future progeny (van der Westhuizen, 2018).

The Ubiquitin-like modifier-activating enzyme gene, *ATG7*, is classified as a Ubiquitin-protein ligase. The molecular functions of this gene include catalytic activity on proteins and ligase activity that

subsequently forms carbon-sulfur bonds (Mi *et al.*, 2019; Pantherdb, 2023b). It is also involved in several biological processes, including, nucleus organisation, proteolysis, protein lipidation, cellular response to starvation, protein modification, mitochondrion autophagy, C-terminal amino acid modification, and lysosomal microautophagy of a cell's nucleus (Mi *et al.*, 2019; Pantherdb, 2023b; Ouellet *et al.*, 2021). Two studies, one focused on Nellore cattle (Silva *et al.*, 2022) and another focused on Japanese Black cattle (Nakanishi *et al.*, 2019), reported *ATG7* to be associated with body weight and skeletal muscle growth during the fattening period. This is due to the important role that autophagy plays in preserving muscle mass under stress conditions and maintaining myofiber integrity (Nakanishi *et al.*, 2019; Silva *et al.*, 2022). The detection of this gene in the EC cows may be explained by the emphasis placed on selection for growth traits in the Bonsmara breed to meet specific breed standards set by the Bonsmara breeders' society (Steyn *et al.*, 2014). As well as to promote superior growth in artificial insemination (AI) bull progeny (Bonsmara SA, 2019).

Free State

There were five genes identified in the FS Bonsmara cow population that have been confirmed cattle, but no associations were found for these genes with growth and adaptation. The genes were *TMEM132B*, *KRAB domain-containing protein*, *SPTB*, *DOCK1*, and *MYO3B*. The transmembrane protein 132B, *TMEM132B*, gene did not fall into any specific Panther molecular function category. It enables various molecular processes in humans, and cattle are an ortholog of the gene (NCBI, 2023b). The *KRAB domain-containing protein*, ZNF747-related, gene, was not placed in any specific Panther molecular function category, but it is classified as a C_2H_2 Zinc finger transcription factor. This gene is involved in nucleolus maintenance, cell proliferation and differentiation, apoptosis, and transcriptional repression of RNA polymerase promoters (Urrutia, 2003; Mi *et al.*, 2019; Pantherdb, 2023n). *SPTB*, spectrin beta chain gene, erythrocytic, fell into the "binding" Panther molecular function category, and is involved in actin filament binding and actin cytoskeleton organisation (Mi *et al.*, 2019; Pantherdb, 2023v). Alterations to the *SPTB* gene's expression have previously been linked to embryonic lethality in cattle (Oishi *et al.*, 2006; Rezende *et al.*, 2021). The dedicator of cytokinesis 1, *DOCK1* gene, fell into the "binding" and "catalytic activity" Panther molecular function categories. Its specific functions include GDP, GTP, and small GTPase binding, and it is involved in cell and neuron migration (Mi *et al.*, 2019; Pantherdb, 2023f). Its involvement in neuron migration has associated the gene with brain development and may influence neuronal behaviour and variations in bovine temperament (Valente *et al.*, 2016).

Of the nine genes identified from the FS cows, three were associated with growth and adaptation. These were *GTDC1*, *PRKG1*, and *ADRA1A*. *GTDC1* and *PRKG1* were not placed in any Panther molecular function category, whilst *ADRA1A* was classed into the "molecular transducer activity" function (Mi *et al.*, 2019).

The glycosyltransferase-like domain-containing protein 1, *GTDC1*, is classified as a glycosyltransferase. This gene is involved in glycan chain biosynthesis, especially the glycan classes in the Golgi apparatus (Varki *et al.*, 2022). It has also been reported to be associated with IMF percentage (Bolormaa *et al.*, 2011). IMF plays an important role in the palatability and quality of meat and thus the gene is targeted for selection in Bonsmara cattle (Nguyen *et al.*, 2021; Linde *et al.*, 2023). The *GTDC1* gene is located close to the *HIBCH* gene, which is close to the *MSTN* gene. Thus, the detection of the gene in the FS cows is likely due to the targeted selection of the bovine myostatin gene for development and regulation of skeletal muscle as well as selection for IMF content (Sheng *et al.*, 2021).

The cGMP-dependent protein kinase 1, *PRKG1*, is classified as a non-receptor serine/threonine protein kinase. The broad functions of this gene include the enabling of ATP binding, cGMP binding, cGMP-dependent protein kinase activity, calcium channel regulator activity, and protein serine kinase activity (Mi *et al.*, 2019; Pantherdb, 2023q). It has been reported to be associated with feed conversion efficiency (FCE) and residual feed intake (RFI) (Lonergan *et al.*, 2010; Sherman *et al.*, 2010; Taye *et al.*, 2017). Vajana *et al.* (2018) reported that the gene was similarly identified in a landscape genomics analysis study on Ugandan cattle and their adaptability to East Coast Fever. They reported that due to its inflammatory response, *PRKG1* was also proposed to be associated with tick resistance in South African cattle (Mapholi *et al.*, 2016). The detection of this gene in the FS cows could be explained by one of the province's predominant bioregions – the Dry Highveld grassland (Mucina *et al.*, 2006). Webb *et al.* (2017) reported that Dry Highveld grassland and Eastern Kalahari bushveld bioregions were associated with higher 18MW compared to Central bushveld and Mesic Highveld grassland bioregions. Given the gene's reported association with FCE and RFI (Lonergan *et al.*, 2010; Sherman *et al.*, 2010; Taye *et al.*, 2017), the vegetation in the Dry Highveld grassland could be complementary to these traits, creating favourable conditions for *PRKG1* expression.

The Alpha-1A adrenergic receptor gene, *ADRA1A*, is involved in several biological processes. Some of these include the activation of phospholipase C activity, cAMP-mediated signalling, and the regulation thereof, activation of adenylate cyclase activity and the regulation thereof, inositol phosphate-mediated signalling, and the regulation of calcium ion concentration in cells (Mi *et al.*, 2019; Pantherdb, 2023a). The gene also encodes an α_1-adrenergic receptor for catecholamines (Hromádková *et al.*, 2020). Expression of the *ADRA1A* gene was reported to be higher in the ileum and adrenal glands of colostrum fed calves (Hromádková *et al.*, 2020). Subsequently, it was speculated that the α_1-adrenergic receptors coded for by the gene, may play a role in fundamental digestive function regulation and mucosal immunity (Santulli *et al.*, 2012; Scanzano & Cosentino, 2015; Hromádková *et al.*, 2020). The presence of more α_1-adrenergic receptors due to higher expression of *ADRA1A*, were thought to ultimately aid in the adaptation of the neonate to its new environment outside of the womb, as well as facilitate proper digestive functioning (Hromádková *et al.*, 2020). There were no calves included in the group of FS cows investigated in this study, therefore the detection of this gene in the FS cows may be attributed to

adaptive digestive responses to environmental stresses brought on by the recent droughts (Walker & Drouillard, 2012; Howell *et al.*, 2022).

North-West

The three NW genes that have been confirmed in cattle but were not associated with growth or adaptation included *TSC22D1*, *RTN4RL1*, and *SPTB*. To date, the TSC22 domain family protein 1, *TSC22D1*, has mostly been found to regulate the transcription of multiple genes in humans (GeneCards, 2023b), but in cattle it may be associated with maternal calving difficulty (Purfield *et al.*, 2020). This gene was not placed into a specific Panther molecular function category. The reticulon 4 receptor like 1 gene, *RTN4RL1*, was placed into the "binding" Panther molecular function category and it enables signaling receptor binding and heparin binding. It is also involved in the biological processes of axon guidance (Mi *et al.*, 2019; Pantherdb, 2023y).

Only two of the eight genes identified in the NW Bonsmara cow analyses were associated with growth and adaptation, *KCNJ16* and *CSMD3*. *KCNJ16* was placed into the Panther molecular function "transporter activity" category, while *CSMD3* was not placed in any Panther molecular function category (Mi *et al.*, 2019).

The potassium inwardly rectifying channel subfamily J member 16, *KCNJ16*, is classified as an ion channel and include voltage-gated potassium channel and ligand-gated ion channel activity (Mi *et al.*, 2019; Pantherdb, 2023m). It is also involved in biological processes which include potassium ion transmembrane transport and the regulation thereof (Mi *et al.*, 2019; Pantherdb, 2023m). Sammad *et al.* (2022) reported that the gene may be a participating candidate gene in the homeostatic modifications that take place in response to heat stress. A negative relationship between temperature and growth performance traits was reported by Webb *et al.* (2017). This relationship is observed because of the negative influence that high ambient temperatures have on the energy status of cattle. Of the various summer month temperatures analysed alongside the cows' genomic SNP data in LEA v.3.8.0, the NW summer month temperatures were the highest. This might explain the need for homeostatic response to heat stress, and subsequent expression of the *KCNJ16* gene in cows born and raised in the NW province.

CSMD3 (also detected in the EC cows) is associated with body size and stature in cattle (Ghoreishifar *et al.*, 2020). Heavier 18MW have previously been reported to be strongly associated with Dry Highveld grassland and Eastern Kalahari bushveld (Webb *et al.*, 2017). Both of which are the predominant bioregions in the NW province (Mucina *et al.*, 2006). The association of the Dry Highveld grassland and Eastern Kalahari bushveld bioregions with larger 18MW may be an explanation for the detected *CSMD3* gene in the NW cows. This possible explanation is in addition to the gene's close location, on BTA14, to *LOC781881*, which has been associated with IMF (Bolormaa *et al.*, 2011).

Out of the 25 detected genes, nine were associated with growth performance or adaptation. Only one gene (*CSMD3*) out of those nine growth/adaptation related genes was detected in multiple provinces (EC and NW). The variable expression of growth performance and adaptation related genes between the provinces may be due to different breeders having different breeding objectives. Variation in breeding system, expenses, factors that influence cattle performance, among other considerations, influence an individual breeder's breeding objectives (Kluyts *et al.*, 2003). Additionally, the environmental variables used for analyses were recorded from 2016 to 2021. During this time, South Africa experienced two significant droughts during 2015 to 2016 and 2018 to 2020 (Meza *et al.*, 2021). The variability in the gene expression observed between the provinces may also be attributable to differing degrees of adaptation in response to the periods of drought. The extent to which the droughts affected each of the environments would explain gene suppression or expression differences between the cow populations (Cavalli & Heard, 2019). The detection of at least one growth related gene from all three provinces, however, supports the nature of the Bonsmara stud breeding industry. Where optimal growth performance is targeted in selection to ensure that cattle meet breed standards to qualify for registration and strong selection pressure is applied for growth performance in AI bulls (Bonsmara SA, 2019).

GWAS result comparison

The GWAS analysis for 18MW identified four significant SNPs in all cows from the three provinces, namely Hapmap49016-BTA-110674, Hapmap32099-BTA-151095, and Hapmap59651-rs29009956; and ARS-BFGL-NGS-26337. Of these SNP variants, only one was found to be associated with a classified gene, namely SNP ARS-BFGL-NGS-26337, associated with sphingomyelin synthetase 1, *SGMS1* (Mi *et al.*, 2019; Pantherdb, 2023t). The other three SNPs were all found on BTA6 but their association to pathways or functions was undetermined. The molecular function of the *SGMS1* gene is to transfer phosphorus-containing groups from one compound to another. The biological functions it is involved in include phospholipid biosynthetic processes and ceramide biosynthetic processes (Pantherdb, 2023t). Specific functions of the *SGMS1* gene related to cattle performance have not been reported. However, a study on the genetic variation of RFI in Australian Angus cattle, documented that *SGMS1* was expressed significantly higher in bulls strongly selected for RFI, than in bulls selected for low RFI (de Las Heras-Saldana *et al.*, 2019).

None of the significantly associated SNPs identified from the GWAS with 18MW, were also detected from the LEA v.3.8.0 analysis based on the environmental variables. However, that was not an indication of poor efficacy of the LEA v.3.8.0 analyses, as the LEA v.3.8.0 analysis did detect the nine candidate SNP variant loci that were in some way associated with genes involved with growth performance.

With the results from this study, it can be concluded that there is potential to investigate growth performance using a landscape genomics approach. The LEA v.3.8.0 analysis was successful in detecting candidate loci that were associated with genes previously reported to be involved with growth performance or adaptation. Gene annotation of the detected candidate genes was helpful in gaining understanding of their functions and the processes they are involved in.

Chapter 5: Critical review, general conclusion, and recommendations

5.1 Conclusion

This study was the first attempt to perform a landscape genomic analysis on SA Bonsmara cattle, with the aim to search for genetic and environmental associations with growth performance. It is of interest to have more information on size and genetic adaptation to a specific environment. Genotypes of Bonsmara cows from the Eastern Cape, Free State and North-West provinces of South Africa, were analysed using a landscape genomic approach in the form of LFMM analysis through the LEA v.3.8.0 package in R. Landscape genomics analysis seeks to establish whether ecological-genotypic associations exist for an organism and the environment it originated from and inhabits (Gugger *et al.*, 2021). To investigate such associations effectively with a landscape genomic approach, in this case LFMM via LEA v.3.8.0, the genotypic data should have some historic exposure to the ecological climate. Historic weather data was available from the weather bureau to provide an extensive view of the provinces included. In the study, the LEA v.3.8.0 analysis provided results on SNP-climate associations in the form of candidate loci that were annotated to identify several genes that are all to some extent, influenced by the temperature, relative humidity, and rainfall.

Of the thousands of candidate loci detected from the LEA v.3.8.0 analysis, 60 SNPs were chosen for gene annotation. The annotation determined 25 classified genes where nine, *CSMD3*, *CDH18*, *MAP4K3*, *HIBCH*, *ATG7*, *GTDC1*, *ADRA1A*, *PRKG1*, and *KCNJ16* were associated with growth performance or adaptation. *CSMD3* has been associated with body size and stature in cattle (Ghoreishifar *et al.*, 2020). *CDH18* has been reported to be significantly associated with age at first calving (Ahmad *et al.*, 2023). *MAP4K3* is believed to have a positive influence on growth performance when an animal is supplemented with probiotics (Kulkarni *et al.*, 2022). *HIBCH* has been found near the bovine myostatin gene and in a QTL region defining carcass fat (Sharma *et al.*, 1999; Aliloo, 2020; Kenny *et al.*, 2022). The *ATG7* gene is involved with cellular autophagy which has been reported to play an important role in body weight and skeletal muscle growth in cattle (Nakanishi *et al.*, 2019; Silva *et al.*, 2022). *GTDC1* has been reported to be associated with intramuscular fat percentage (Bolormaa *et al.*, 2011). *ADRA1A* encodes an α_1-adrenergic receptor for catecholamines, which might be involved in the regulation of fundamental digestive functions and glucose metabolism (Hromádková *et al.*, 2020). *PRKG1* has been associated with FCE, RFI, and tick resistance (Lonergan *et al.*, 2010; Sherman *et al.*, 2010; Mapholi *et al.*, 2016; Taye *et al.*, 2017). *KCNJ16* has been reported to play a role in homeostatic processes that take place in response to heat stress (Sammad *et al.*, 2022).

5.2 Efficacy of investigation

A limitation for landscape genomic analyses lies in the nature of the South African cattle stud breeding industry. Cows and bulls are raised on farms but change ownership throughout their lives as they are purchased at different ages and moved between provinces and climatic biomes. This can happen multiple times in an animal's life for the purpose of increasing the gene pool in a breeding herd. Given the established differences in climate and environment between the South African provinces, this also means that the cattle are subsequently exposed to varying climates. This hinders the establishment of heritable local genetic adaptation. Gugger *et al.* (2021) mentioned that the stability of the Californian endemic oak genetics in a 66entralized location over an extended period improved the efficacy of the landscape genomic analysis. Therefore, a landscape genomic approach is most successfully applied when the genetic information has been exposed to a single geographical environment over the course of decades (Storfer *et al.*, 2018; Gugger *et al.*, 2021). This factor was taken into consideration during the animal selection process, however, not to the extent of decades. When effort was made to identify animals that had parents and grandparents born and raised or that spent a significant portion of their lives in the same geographic region as their offspring, not a single cow out of the chosen 765 could be identified. This was due to the parents and grandparents having been moved around the country for stud breeding purposes. The overall effect of this is that the cows used for analysis did not have inherited genetic information that had been exposed to a single environment for decades, meaning that the adaptation they express to their environments is more likely to be temporary instead of permanently inherited from their parents.

The environmental variables chosen for analysis were summer month temperatures, winter month temperatures, relative humidity, and average annual precipitation (Gugger *et al.*, 2021; Shryock *et al.*, 2021). The environmental data received from the SA Weather Bureau was extensive, dating back to 2001 for most of the weather stations. Ideally, more historic weather data is desirable for landscape genomic analyses as it would correspond to historic genetic data of the species being investigated and would consider weather trends over long time periods in the landscape genomic analysis (Gugger *et al.*, 2021). However, there were gaps in the raw weather data received for this study, which limited the available climatic variable data for analysis to the short period of 2016 to 2021. More extensive and complete data would have improved the accuracy of the landscape genomics analysis.

The results obtained from the LEA v.3.8.0 analysis indicated that there was some association between the environment and cow growth performance. Nine out of the 25 classified-gene-associated SNP variants were documented to have some association with growth performance or adaptation. The study confirmed the need for follow up analyses with larger numbers of genotypes. Whole-genome sequencing (WGS) data could provide more dense genomic information for an association study.

5.3 Recommendations

There are several modifications that can be made to the methodology to potentially improve the outcome of the results. The first would be to obtain more extensive and complete weather data for analysis. WorldClim weather data can be downloaded using the raster package in R. This data would be more extensive and completed data records from the early 2000s can also be accessed (Hijmans *et al.*, 2005; Mdladla *et al.*, 2018).

Another avenue that could be investigated is to identify specific SNPs on the panels that are known to have an association with growth traits or adaptive traits and to test the animals' genotypes for these specific SNPs against the environmental variables (Joost *et al.*, 2007). This would be an alternative to the genome-wide investigation conducted for this study. Analysing a small number of specific SNPs would allow for a more focused approach at investigating the existence of SNP-climate associations. Additionally, to achieve alternate associations for growth performance traits, the GWAS could have been conducted using de-regressed EBVs instead of phenotypes.

The use of WGS data will offer a more expansive scope of genomic information for investigation of GEA (Yoder *et al.*, 2014; Rellstab *et al.*, 2015). Analysing a larger dataset of genotypes, instead of 25 272, might allow for the detection of more SNP-climate associations than were detected in this study. Bhardwaj *et al.* (2023) obtained the genetic information for the cattle they investigated from a 777k BovineHD BeadChip. This allowed for the identification of 1305 significant SNPs. Yoder *et al.* (2014) utilised a WGS approach to identify a data set of nearly two million SNPs for their environmental association study. Along with denser genomic information for analysis, Stucki *et al.* (2017) suggested the use of the SAMβADA approach as it allows for the analysis of large WGS datasets on desktop computers. It has also been reported to process data ten times faster than a LFMM approach (Stucki *et al.*, 2017).

In this first attempt to use landscape genomics it was demonstrated that it can be used for identifying candidate loci. Follow up studies should however use larger datasets, with more historic weather data and genotypes from whole genome sequencing.

References

©2003-2023 Tableau Software, LLC, A Salesforce Company. ALL. Accessed: August 2021. https://www.tableau.com/products/desktop

Abin, S., Theron, H.E. & van Marle-Köster, E., 2016. Population structure and genetic trends for indigenous African beef cattle breeds in South Africa. S. Afr. J. Anim. Sci. 46(2), 152-156.

Adjei-Fremah, S., Ekwemalor, K., Asiamah, E.K., Ismail, H., Ibrahim, S. & Worku, M., 2018. Effect of probiotic supplementation on growth and global gene expression in dairy cows. J. Appl. Anim. Res. 46(1), 257-263.

Ahmad, S.F., Singh, A., Gangwar, M., Kumar, S., Dutt, T. & Kumar, A., 2023. Haplotype-based association study for production and reproduction traits in multigenerational Vrindavani population. Gene. 867(147365).

Akanno, E.C., Chen, L., Abo-Ismail, M.K., Crowley, J.J., Wang, Z., Li, C., Basarab, J.A., MacNeil,

M.D. & Plastow, G.S., 2018. Genome-wide association scan for heterotic quantitative trait loci in multi-breed and crossbred beef cattle. Genet. Sel. Evol. 50, 1-13.

Alexander, D. H., Novembre, J. & Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19, 1655–1664.

Alexander, D. H., Shringarpure, S. S., Novembre, J. & Lange, K., 2020. ADMIXTURE 1.3 Software Manual.

Aliloo, H., Mrode, R., Okeyo, A.M. & Gibson, J.P., 2020. Ancestral haplotype mapping for GWAS and detection of signatures of selection in admixed dairy cattle of Kenya. Front. Genet. 11, 544.

Alves, B.G., Alves, K.A., Lúcio, A.C., Martins, M.C., Silva, T.H., Alves, B.G., Braga, L.S., Silva, T.V., Viu, M.A., Beletti, M.E., Jacomini, J.O., Santos, R.M. & Gambarini, M.L., 2014. Ovarian activity and oocyte quality associated with the biochemical profile of serum and follicular fluid from Girolando dairy cows postpartum. Anim. Reprod, Sci. 146(3-4). 117-25.

Amundson, J.L., Mader, T.L., Rasby, R.J. & Hu, Q.S. 2006. Environmental effects on pregnancy rate in beef cattle. J. Anim. Sci. 84, 3415-3420.

Andrews, K. R, Good, J. M, Miller, M. R., Luikart, G. & Hohenlohe, P. A., 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat. Rev. Genet. 17, 81–92.

Angel, S.P., Amitha, J.P., Rashamol, V.P., Vandana, G.D., Savitha, S.T., Afsal, A., Bagath, M., Krishnan, G. & Sejian, V., 2018. Climate Change and Cattle Production-Impact and Adaptation. J. Vet. Med. Res. 5(4), 1134.

Barbosa, F.C., Oliveira Junior, W., Faria, J.G.K., Vieira, D.S., Guimarães, E.C. & Mundim, A.V., 2022. Influence of age and sex on blood biochemical profile of Bonsmara cattle breed up to two years. Ciência Animal Brasileira. 23.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A. & Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data. 5(1), 1-12.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B. Met. 57(1), 289-300.

Ben-Jemaa, S., Mastrangelo, S., Lee, S.H., Lee, J.H. & Boussaha, M., 2020. Genome-wide scan for selection signatures reveals novel insights into the adaptive capacity in local North African cattle. Sci. Rep-UK. 10(1), 1-14.

Bermingham, M.L., Bishop, S.C., Woolliams, J.A., Pong-Wong, R., Allen, A.R., McBride, S.H., Ryder, J.J., Wright, D.M., Skuce, R.A., McDowell, S.W. & Glass, E.J., 2014. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. Heredity. 112(5), 543-551.

Bernabuccil, U., Lacetera, N., Baumgard, L.H., Rhoads, R.P., Ronchi, B. & Nardone, A., 2010. Metabolic and hormonal acclimation to heat stress in domesticated ruminants. Animal. 4, 1167-1183.

Berry, D.P., Conroy, S., Pabiou, T. & Cromie, A.R., 2017. Animal breeding strategies can improve meat quality attributes within entire populations. Meat Sci. 132, 6-18.

Bewley, J.M., Grott, M.W., Einstein, M.E. & Schutz, M.M., 2008. Impact of intake water temperatures on reticular temperatures of lactating dairy cows. J. Dairy Sci. 91(10), 3880-3887.

Bhardwaj, S., Singh, S., Ganguly, I., Bhatia, A.K. & Dixit, S.P., 2023. Deciphering local adaptation of native Indian cattle (Bos indicus) breeds using landscape genomics and in-silico prediction of deleterious SNP effects on protein structure and function. 3 Biotech. 13(3).

Blamey, R.C., Ramos, A.M., Trigo, R.M., Tomé, R. & Reason, C.J.C., 2018. The influence of atmospheric rivers over the South Atlantic on winter rainfall in South Africa. J. Hydrometeorol. 19(1), 127-142.

Bolormaa, S., Neto, L.P., Zhang, Y.D., Bunch, R.J., Harrison, B.E., Goddard, M.E. & Barendse, W., 2011. A genome-wide association study of meat and carcass traits in Australian cattle. J. Anim. Sci. 89(9), 2297-2309.

Bolormaa, S., Pryce, J.E., Reverter, A., Zhang, Y., Barendse, W., Kemper, K., Tier, B., Savin, K., Hayes, B.J. & Goddard, M.E., 2014. A Multi-Trait, Meta-Analysis for Detecting Pleiotropic Polymorphisms for Stature, Fatness and Reproduction in Beef Cattle. Plos. Genet. 10, 1-23.

Bonsma, J.C., 1980. Livestock production-a global approach. Tafelburg Publishers Ltd., Cape Town, South Africa.

Bonsma, J.C., 1983. Man must measure: Livestock Production. Agi Books, Cody, W.Y., USA. Bonsma, J.C. & Bonsma, J., 1985. The Bonsmara Beef Cattle Breed. Bonsmara Cattle Breeders' Society 21st Anniversary Publication. Alpha Bonsma Group, Botswana.

Bonsmara SA, 2019. Codes for rejection reasons and Minimum breed standards. Schedule B. Chapter 3. Growth. Accessed April 2023. https://bonsmara.co.za/wp-content/uploads/2019/03/ENGELS-BREED-STANDARDS.pdf

Bosman, L., van Marle-Köster, E., Van der Westhuizen, R.R., Visser, C. & Berry, D.P., 2017. Population structure of the South African Bonsmara beef breed using high density single nucleotide polymorphism genotypes. Livest. Sci. 197, 102-105.

Bourdon, R.M. & Brinks, J.S., 1982. Genetic, environmental and phenotypic relationships among gestation length, birth weight, growth traits and age at first calving in beef cattle. J. Anim. Sci. 55(3), 543-553.

Brand, T.S., Terblanche, I., Van der Merwe, D.A., Jordaan, J.W. & Dreyer, O., 2021. Predicting the growth and feed intake of Bonsmara cattle on forage based rearing systems. Preprint from Research Square.

Brand-Williams, W. & Hayes, H., 2020. A long experience in genetic evaluation. https://www6.jouy.inrae.fr/gabi_eng/Our-Research/History-of-our-research-axes/Genetic- evaluations.

Bridges, P.J., Brusie, M.A. & Fortune, J.E., 2005. Elevated temperature (heat stress) in vitro reduces androstenedione and estradiol and increases progesterone secretion by follicular cells from bovine dominant follicles. Domest. Anim. Endocrin. 29, 508–522.

Brody, S. & Lardy, H.A., 1946. Bioenergetics and growth. J. Phys. Chem. 50(2), 168-169.

Buchmann, R. & Hazelhurst, S., 2014. Genesis Windows10 Java8 0.3.0 210808. University of Witwatersrand, Johannesburg.

Bureau for Food and Agricultural Policy (BFAP), 2021. Outlook for animal products meat, eggs, and wool. 65-78. Accessed February 2023.

Burrow, H.M., 2001. Variances and covariances between productive and adaptive traits and temperament in a composite breed of tropical beef cattle. Livest. Prod. Sci. 70(3), 213-233.

Bush, W.S. & Moore, J.H., 2012. Genome-wide association studies. Plos. Comput. Biol. 8, e1002822.

Buskirk, D.D., Faulkner, D.B. & Ireland, F.A., 1995. Increased postweaning gain of beef heifers enhances fertility and milk production. J. Anim. Sci. 73(4), 937-946.

Buzanskas, M.E., Grossi, D.A., Ventura, R.V., Schenkel, F.S., Sargolzaei, M., Meirelles, S.L., Mokry, F.B., Higa, R.H., Mudadu, M.A., da Silva, M.V.B. & Niciura, S.C., 2014. Genome-wide association for growth traits in Canchim beef cattle. Plos One. 9(4), e94802.

Carraud, A., 1990. Probiotics in beef cattle. Bulletin des GTV. 6, 65-73.

Carnegie, G.K., Means, C.K. & Scott, J.D., 2009. A-kinase anchoring proteins: from protein complexes to physiology and disease. IUBMB Life. 61(4), 394-406.

Cavalli, G. & Heard, E., 2019. Advances in epigenetics link genetics to the environment and disease. Nature. 571(7766), 489-499.

Chadyiwa, M.C. & Wepener, M., 2021. Beef Production. ARC-ISCW Climate-Smart Agriculture Training Module 8.

Chen, M., Pan, D., Ren, H., Fu, J., Li, J., Su, G., Wang, A., Jiang, L., Zhang, Q. & Liu, J.F., 2016. Identification of selective sweeps reveals divergent selection between Chinese Holstein and Simmental cattle populations. Genet. Sel. Evol. 48(1), 1-12.

Cheng, M., McCarl, B. & Fei, C., 2022. Climate change and livestock production: a literature review. Atmosphere-Basel. 13(1), 140.

Chikosi, E.S., Mugambiwa, S.S., Tirivangasi, H.M. & Rankoana, S.A., 2019. Climate change and variability perceptions in Ga-Dikgale community in Limpopo Province, South Africa. Int. J. Clim. Chang. Str. 11(3), 392-405.

Climate: Mpumalanga. Climate-data.org. Accessed February 2023. https://en.climate-data.org/africa/south-africa/kwazulu-natal/mpumalanga-27228/

Climate: Northern Cape. Climate-data.org Accessed February 2023. https://en.climate- data.org/africa/south-africa/northern-cape-470/

Climate in Eastern Cape South Africa. Worlddata.info. Accessed February 2023. https://www.worlddata.info/africa/south-africa/climate-eastern-cape.php#:~:text=Eastern%20Cape%20is%20one%20of,bathe%20in%20the%20warm%20season

Climate in Gauteng South Africa. Worlddata.info. Accessed February 2023. https://www.worlddata.info/africa/south-africa/climate-gauteng.php#:~:text=Gauteng%20is%20one%20of%20the,sometimes%20up%20to%2028%20degrees

Climate in KwaZulu-Natal South Africa. Worlddata.info. Accessed February 2023. https://www.worlddata.info/africa/south-africa/climate-kwazulu-natal.php#:~:text=The%20climate%20is%20almost%20moderate,is%20from%20May%20to%20August

Climate in Limpopo South Africa. Worlddata.info. Accessed February 2023. https://www.worlddata.info/africa/south-africa/climate-limpopo.php

Climate in Mpumalanga South Africa. Worlddata.info. Accessed February 2023. https://www.worlddata.info/africa/south-africa/climate-mpumalanga.php

Climate in North-West South Africa. Worlddata.info. Accessed February 2023. https://www.worlddata.info/africa/south-africa/climate-north-west.php

Climate in Orange Free State South Africa. Worlddata.info. Accessed February 2023. https://www.worlddata.info/africa/south-africa/climate-orange-free-state.php

Collier, R.J. & Zimbelman, R.B. 2007. Heat stress effects on cattle: What we know and what we don't know. 22nd Annual Southwest Nutrition & Management Conference, Tempe, AZ. 76.

Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J.K., 2010. Using environmental correlations to identify loci underlying local adaptation. Genetics. 185(4), 1411-1423.

Cortellari, M., Barbato, M., Talenti, A., Bionda, A., Carta, A., Ciampolini, R., Ciani, E., Crisà, A., Frattini, S., Lasagna, E. & Marletta, D., 2021. The climatic and genetic heritage of Italian goat breeds with genomic SNP data. Sci. Rep-UK. 11(1), 1-12.

Crook, B.J., Neser, F.W.C. & Bradfield, M.J., 2010. Genetic parameters for cow weight at calving and at calf weaning in South African Simmental cattle. S. Afr. J. Anim. Sci. 40, 140-144.

da Gama, M.P.M., de Paz, C.C.P., Junior, M.L.S., Ventura, H.T., Bernardes, P.A. & El Faro, L., 2021. Genetic associations between growth traits and age at first calving in Guzera cattle using random regression models. Livest. Sci. 250, 104585.

Daramola, J.O., Abioja, M.O. & Onagbesan, O.M., 2012. In: Heat stress impact on livestock production. Environmental stress and amelioration in livestock production. Sijan, V., Naqvi, S.M.K., Ezeji, T., Lakritz, J., and Lal, R. (5), 53-73. Springer Berlin Heidelberg.

Dash, S., Chakravarty, A.K., Singh, A., Upadhyay, A., Singh, M. & Yousuf, S., 2016. Effect of heat stress on reproductive performances of dairy cattle and buffaloes: A review. Vet World. 9(3). 235-244.

De Jager, J.M., 1993. Geographic and agro-ecological features of Southern Africa. In: Livestock Production systems Principles and Practice. Ed. Maree, C. & Casey, N.H. Pretoria: Agri Development Foundation, South Africa.

de Las Heras-Saldana, S., Clark, S.A., Duijvesteijn, N., Gondro, C., van der Werf, J.H. & Chen, Y., 2019. Combining information from genome-wide association and multi-tissue gene expression studies to elucidate factors underlying genetic variation for residual feed intake in Australian Angus cattle. BMC genomics. 20(1), 1-16.

De Vos, J., 2018. A Genome wide association study of carcass traits based on Real Time Ultrasound in South African Nguni cattle (Doctoral dissertation, University of Pretoria).

Denson, E., Wasko, C. & Peel, M.C., 2021. Decreases in relative humidity across Australia. Environ. Res. Lett. 16(7), 074023.

Dhama, K., Mahendran, M., Tomar, S. & Chauhan, R.S., 2008. Beneficial effects of probiotics and prebiotics in livestock and poultry: the current perspectives. Intas Polivet. 9(1), 1-12.

Dickerson, G.E., 1970. Efficiency of Animal Production—Moulding the Biological components. J. Anim. Sci. 30, 849-859.

Dickerson, G.E. 1978. Animal size and efficiency: basic concepts. Anim. Prod. 27, 367-379.

Du Preez, J.H., Giesecke, W.H., Barnard, M.L., Erasmus, A.G., Hatting, P.J., Eisenburg, B.E., Willemse, J.J.C. & Kruger, R. 1992. Heat stress on dairy cattle and other livestock under South African conditions. Tech. comm. Dept. Agric. Dev. Repub. S. Afr.

Dzavo, T., Zindove, T.J., Dhliwayo, M. & Chimonyo, M., 2019. Effects of drought on cattle production in sub-tropical environments. Trop. Anim. Health Pro. 51, 669-675.

Ellery, W.N., Scholes, R.J. & Scholes, M.C., 1995. The distribution of sweetveld and sourveld in South Africa's grassland biome in relation to environmental factors. Afr. J. Range. For. Sci. 12(1), 38- 45.

Elum, Z.A., Modise, D.M. & Marr, A., 2017. Farmer's perception of climate change and responsive strategies in three selected provinces of South Africa. Climate Risk Management. 16, 246-257.

Engelbrecht, C.J., Landman, W.A., Engelbrecht, F.A. & Malherbe, J., 2015. A synoptic decomposition of rainfall over the Cape south coast of South Africa. Clim. Dynam. 44, 2589-2607.

European Variation Archive(EVA). EMBL-EBI. Accessed: 27 March 2023. https://www.ebi.ac.uk/eva/?Variant-Browser&species.

Fariello, M.I., Boitard, S., Naya, H., SanCristobal, M. & Servin, B., 2013. Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. Genetics. 193(3), 929-941.

Flori, L., Moazami-Goudarzi, K., Alary, V., Araba, A., Boujenane, I., Boushaba, N., Casabianca, F., Casu, S., Ciampolini, R., Coeur D'Acier, A. & Coquelle, C., 2019. A genomic map of climate adaptation in Mediterranean cattle breeds. Mol. Ecol. 28(5), 009-1029.

Food and Agriculture Organization of the United Nations, 2016. Drought characteristics and management in the Caribbean. Retrieved from http://www.fao.org/3/a-i5695e.pdf

Forester, B. R., Lasky, J. R., Wagner, H. H., and Urban, D. L., 2018. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Mol. Ecol. 27(9), 2215–2233.

Foster, L.A., Fourie, P.J. & Neser, F.W.C., 2009. Effect of heat stress on six beef breeds in the Zastron district: The significance of breed, coat colour and coat type. S. Afr. J. Anim. Sci. 39(sup-1), 224-228.

Freitas, P.H., Wang, Y., Yan, P., Oliveira, H.R., Schenkel, F.S., Zhang, Y., Xu, Q. & Brito, L.F., 2021. Genetic diversity and signatures of selection for thermal stress in cattle and other two bos species adapted to divergent climatic conditions. Front. Genet. 12, 604823.

Frichot, E. & François, O., 2015a. LEA: An R package for landscape and ecological association studies. Methods Ecol. Evol. 6(8), 925–929.

Frichot, E. & François, O., 2015b. A short manual for LFMM version 1.4 (command line version). Retrieved from http://membres-timc.imag.fr/ Olivier.Francois/lfmm/files/note.pdf.

Frichot, E., Schoville, S. D., Bouchard, G. & François, O., 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. Mol. Biol. Evol. 30(7), 1687–1699.

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., and Francois, O., 2014. Fast and efficient estimation of individual ancestry coefficients. Genetics. 196, 973–983.

Furstenburg, D. & Scholtz, M.M., 2009. Global climate change and animal production in southern Africa: A short review. Submitted: Livest. Sci. (Supplement-10th Wrld. Cong. Anim. Prod.)

Garrick, D.J., 2011. The nature, scope and impact of genomic prediction in beef cattle in the United States. Genet. Sel. Evol. 43(1), 1-11.

Garrick, D.J. & Enns, R.M., 2003. How best to achieve genetic change. In: Proc. Beef Imp. Fed. 35th Annu. Res. Symp. & Annu. Mtg., Lexington, Kentucky. 28-31.

Gathura, D.M., Muasya, T.K. & Kahi, A.K., 2020. Meta-analysis of genetic parameters for traits of economic importance for beef cattle in the tropics. Livest. Sci. 242, 104306.

Gaughan, J. & Cawdell-Smith, A.J., 2015. Impact of climate change on livestock production and reproduction. In: Climate change Impact on livestock: adaptation and mitigation. Johnson, J.S., Abuajamieh, M., Fernandez, M.S., Siebert, J.T., Stoakes, S.K., Nteeba, J., Keating, A.F., Ross, J.W., Rhoads, R.P. and Baumgard, L. 51-60. New Delhi: Springer India.

Gaughan, J.B., Sejian, V., Mader, T.L. & Dunshea, F.R., 2019. Adaptation strategies: ruminants. Anim. Front. 9(1), 47-53.

GeneCards. GLG1 gene. Entrez Gene Summary provided by Alliance of Genome Rsesources, April 2022. Accessed: 7 April, 2023a. https://www.genecards.org/cgi-bin/carddisp.pl?gene=GLG1.

GeneCards. TSC22D1 gene. Entrez Gene Summary provided by RefSeq, August 2011. Accessed: 7 April, 2023b. https://www.genecards.org/cgi-bin/carddisp.pl?gene=TSC22D1.

GeneCards. WDR25 gene. Entrez Gene Summary provided by RefSeq, April 2017. Accessed: 7 April, 2023c. https://www.genecards.org/cgi-bin/carddisp.pl?gene=WDR25.

Georges, M., Charlier, C. & Hayes, B., 2019. Harnessing genomic information for livestock improvement. Nat. Rev. Genet. 20(3), 135-156.

Ghoreishifar, S.M., Eriksson, S., Johansson, A.M., Khansefid, M., Moghaddaszadeh-Ahrabi, S., Parna, N., Davoudi, P. & Javanmard, A., 2020. Signatures of selection reveal candidate genes involved in economic traits and cold acclimation in five Swedish cattle breeds. Genet. Sel. Evol. 52, 1-15.

Goddard, M.E., & Hayes, B.J., 2011. Genome-wide Association Studies and Linkage Disequilibrium in cattle. Bovine Genomics, first edition. Ch13, 192-210.

Goitom, S., Gicheha, M.G., Njonge, F.K. & Kiplangat, N., 2021. Landscape genomics and selection signatures of local adaptation of Eritrean indigenous cattle along environmental gradients. Trop. Anim. Health Pro. 53, 1-8. t

Gonzalez-Rivas, P.A., Chauhan, S.S., Ha, M., Fegan, N., Dunshea, F.R. & Warner, R.D., 2020. Effects of heat stress on animal physiology, metabolism, and meat quality: A review. Meat Sci. 162, 108025.

Gouws, A., 2020. Bonsmaras gain a foothold in Botswana. Die Rooi Ras. 4(1), 17.

Govender, K., 2019. Beef Production. KZN Agriculture & Rural Development. https://www.kzndard.gov.za/2-uncategorised/374-beef-production.

Grummer, J.A., Beheregaray, L.B., Bernatchez, L., Hand, B.K., Luikart, G., Narum, S.R. & Taylor, E.B., 2019. Aquatic landscape genomics and environmental effects on genetic variation. Trends Ecol. Evol. 34(7), 641-654.

Gugger, P.F., Fitz-Gibbon, S.T., Albarrán-Lara, A., Wright, J.W. & Sork, V.L., 2021. Landscape genomics of Quercus lobata reveals genes involved in local climate adaptation at multiple spatial scales. Mol. Ecol. 30, 406–423.

Günther, T. & Coop, G., 2013. Robust identification of local adaptation from allele frequencies. Genet. 195(1), 205-220.

Gurgul, A., Semik, E., Pawlina, K., Szmatoła, T., Jasielczuk, I. & Bugno-Poniewierska, M., 2014. The application of genome-wide SNP genotyping methods in studies on livestock genomes. J. Appl. Genet. 55(2), 197-208.

Gwazdauskas, F.C., 1985. Effects of climate on reproduction in cattle. J. Dairy Sci. 68(6), 1568-1578.

Hafez, E.S.E. 1968. Principles of animal adaptation. In: Adaptation of domestic animals. Eds. Hafez, E.S.E. Philadelphia: Lea & Febiger.

Hamersky, S., 2016. Tableau Desktop. Mathematics and Computer Education; Old Bethpage. 50(2), 148-151.

Harvey, S.H., Krien, M.J. and O'Connell, M.J., 2002. Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases. Genome Biol. 3(2), 1-5.

Hayes, B.J., Chamberlain, A.J., Maceachern, S., Savin, K., McPartlan, H., MacLeod, I., Sethuraman, L. & Goddard, M.E., 2009. A genome map of divergent artificial selection between Bos taurus dairy cattle and Bos taurus beef cattle. Anim. Genet. 40(2), 176-184.

Hayes, B.J. & Goddard, M., 2010. Genome-wide association and genomic selection in animal breeding. Genome. 53, 876-883.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A., 2005. Very high-resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25(15), 1965-1978.

Hirschhorn, J.N. & Daly, M.J., 2005. Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. 6, 95.

Horowitz, M. 2001. Heat acclimation: phenotypic plasticity and cues to the underlying molecular mechanisms. J. Therm. Biol. 26, 357-363.

Howell, A., Arsic, N., Brownlie, R. & Griebel, P., 2022. Adrenergic receptor gene expression in bovine leukocytes. Dev. Comp. Immunol. 127, 104271.

Hromádková, J., Suzuki, Y., Pletts, S., Pyo, J., Ma, T., Chen, Y. & Steele, M.A., 2020. Effect of colostrum feeding strategies on the expression of neuroendocrine genes and active gut mucosa-attached bacterial populations in neonatal calves. J. Dairy Sci. 103(9), 8629-8642.

Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P., Hanson, C.E., 2007. Climate change 2007: Impacts, adaptation and vulnerability, working group II contribution to the IPCC fourth assessment report. IPCC.

Alexander, L., Allen, S. and Bindoff, N.L., 2013. Working group I contribution to the IPCC fifth assessment report climate change 2013: The physical science basis summary for policymakers (No. Bajados de Internet/2013). IPCC.

Irwin, S. & Good, D., 2012. Do warm winters tell us anything about summer temperatures and corn yields?. Farmdoc Daily. 2(46).

James, J.W. 2009. Genotype by environment interactions in farm animals: Evolutionary and breeding perspectives on genetic resource management. In: Adaptation and fitness in animal populations Eds. Van der Werf, J., Graser, H., Frankham, R., and Gondro, C. Springer Science + Business Media B.V.

Joost, S., Bonin, A., Bruford, M. W., Després L., Conord, C., Erhardt, G. & Taberlet, P., 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Mol. Ecol. 16, 3955–3969.

Joost, S. & Negrini, R., 2010. Early stirrings of landscape genomics: awaiting next-next generation sequencing platforms before take-off. In: Sustainable Improvement of Animal Production and Health. Odongo, N.E., Garcia, M., and Viljoen, G.J. Food and Agriculture Organization of the United Nations, Rome. 185-189.

Jordaan, F.J., Neser, F.W., Maiwashe, A., King, Z. & Scholtz, M.M., 2021. The environmental impact of changes in cow productivity and its component traits in South Africa's Landrace beef breeds. Front. Anim. Sci. 2, 743229.

Jury, M.R., 2013. Climate trends in southern Africa. S. Afr. J. Sci. 109(1), 1-11.

Kassambara, A., 2023. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.2, https://rpkgs.datanovia.com/rstatix/.

Kattnig, R.M., Winder, J.A., Wallace, J.D. & Bailey, C.C., 1993. Evaluation of biological efficiency of free grazing beef cows under semidesert conditions. J. Anim. Sci. 71, 2601-2607.

Kelsey, A.J. & Colpoys, J.D., 2018. Effects of dietary probiotics on beef cattle performance and stress. J. Vet. Behav. 27, 8-14.

Kemper, K.E., Saxton, S.J., Bolormaa, S., Hayes, B.J. & Goddard, M.E., 2014. Selection for complex traits leaves little or no classic signatures of selection. BMC Genomics. 15(1), 1-14.

Kenny, D., Sleator, R.D., Murphy, C.P., Evans, R.D. & Berry, D.P., 2022. Detection of Genomic Imprinting for Carcass Traits in Cattle Using Imputed High-Density Genotype Data. Front. Genet. 13, 951087.

Kimaro, E.G., Mor, S.M. & Toribio, J.A.L., 2018. Climate change perception and impacts on cattle production in pastoral communities of northern Tanzania. Pastoralism. 8, 1-16.

Kizilkaya, K., Tait, R.G., Garrick, D.J., Fernando, R.L. & Reecy, J.M., 2013. Genome-wide association study of infectious bovine keratoconjunctivitis in Angus cattle. BMC Genet. 14(23).

Kluyts, J.F., Neser, F.W.C. & Bradfield, M.J., 2003. Development of breeding objectives for beef cattle breeding: Derivation of economic values. S. Afr. J. Anim. Sci. 33(3), 142-158.

Konapala, G., Mishra, A.K., Wada, Y. & Mann, M.E., 2020. Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. Nat. Commun. 11(1), 3044.

Korte, A. & Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods. 9, 29-38.

Krebhiel, B. C., Thomas, M. G., Wilson, C. S., Speidel, S. E., Enns, R. M., Paiva, S. R. & Blackburn, H. D., 2019. Evaluation of genetic structure across U.S. climate zones using prominent AI sires of Red Angus cattle. Livest. Sci. 225, 26-31.

Kulkarni, N.A., Chethan, H.S., Srivastava, R. & Gabbur, A.B., 2022. Role of probiotics in ruminant nutrition as natural modulators of health and productivity of animals in tropical countries: an overview. Trop. Anim. Health Pro. 54(2), 110.

Lees, A.M., Sejian, V., Wallage, A.L., Steel, C.C., Mader, T.L., Lees, J.C. & Gaughan, J.B., 2019. The impact of heat load on cattle. Animals. 9(6), 322.

Li, Y., Zhang, X., Mao, R., Yang, J., Miao, C., Li, Z. & Qiu, Y., 2017. Ten Years of Landscape Genomics: Challenges and Opportunities. Front. Plant Sci. 8(2136).

Linde, D.A., van Marle-Köster, E., Scholtz, M.M., Gonda, M.G., Gonzalez-Hernandez, J.L. & MacNeil, M.D., 2023. Differential gene expression in the Longissimus dorsi of Nguni and Bonsmara bulls finished on low and high energy diets. S. Afr. J. Anim. Sci. 53(1), 101-108.

Lipták, T., 1958. On the combination of independent tests. Magyar Tud Akad Mat Kutato Int Kozl. 3, 171-197.

Long, J.A., 2008. Reproductive biotechnology and gene mapping: tools for conserving rare breeds of livestock. Reprod. Domest. Anim. 43, 83-88.

Lonergan, E.H., Zhang, W. & Lonergan, S.M., 2010. Biochemistry of post-mortem muscle—Lessons on mechanisms of meat tenderization. Meat Sci. 86(1), 184-195.

Lotterhos, K.E. and Whitlock, M.C., 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Mol. Ecol. 24(5), 1031-1046.

Lowry, D.B., Hoban, S., Kelley, J.L., Lotterhos, K.E., Reed, L.K., Antolin, M.F. & Storfer, A., 2017. Responsible RAD: Striving for best practices in population genomic studies of adaptation. Mol. Ecol. Resour. 17(3), 366-369.

Maas, J., 1987. Relationship between nutrition and reproduction in beef cattle. Veterinary clinics of North America: food animal practice. 3(3), 633-646.

Magita, S.Y. & Sangeda, A.Z., 2017. Effects of climate stress to pastoral communities in Tanzania: A case of Mvomero District. Livestock Research for Rural Development. 29(160).

Mahlalela, P.T., Blamey, R.C., Hart, N.C.G. & Reason, C.J.C., 2020. Drought in the Eastern Cape region of South Africa and trends in rainfall characteristics. Clim. Dynam. 55, 2743-3759.

Maiwashe, A.N., Bradfield, M.J., Theron, H.E. & Van Wyk, J.B., 2002. Genetic parameter estimates for body measurements and growth traits in South African Bonsmara cattle. Livest. Prod. Sci. 75(3), 293-300.

Makina, S.O., Muchadeyi, F.C., van Marle-Köster, E., Taylor, J.F., Makgahlela, M.L. & Maiwashe, A., 2015. Genome-wide scan for selection signatures in six cattle breeds in South Africa. Genet. Sel. Evol. 47(1), 1-14.

Makina, S.O., Whitacre, L.K., Decker, J.E., Taylor, J.F., MacNeil, M.D., Scholtz, M.M., van Marle-Köster, E., Muchadeyi, F.C., Makgahlela, M.L. & Maiwashe, A., 2016. Insight into the genetic composition of South African Sanga cattle using SNP data from cattle breeds worldwide. Genet. Sel. Evol. 48, 1-7.

Mapholi, N.O., Maiwashe, A., Matika, O., Riggio, V., Bishop, S.C., MacNeil, M.D., Banga, C., Taylor, J.F. & Dzama, K., 2016. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick-Borne Dis. 7(3), 487-497.

Mapiye, C., Chimonyo, M. & Dzama, K., 2009. Seasonal dynamics, production potential and efficiency of cattle in the sweet and sour communal rangelands in South Africa. J. Arid Environ. 73(4-5), 529-536.

Mathevon, M., Buhr, M.M. & Dekkers, J.C.M., 1998. Environmental, management, and genetic factors affecting semen production in Holstein bulls. J. Dairy Sci. 81(12), 3321-3330.

Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M.P., O'Connell, J., Moore, S.S., Smith, T.P., Sonstegard, T.S. & Van Tassell, C.P., 2009. Development and characterization of a high-density SNP genotyping assay for cattle. Plos One. 4, e5350.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R., Thormann, A., Flicek, P. & Cunningham, F., 2016. The ensembl variant effect predictor. Genome Biol. 17(1), 1-14.

Mdladla, K., Dzomba, E. F. & Muchadeyi, F. C., 2017. The potential of landscape genomics approach in the characterization of adaptive genetic diversity in indigenous goat genetic resources: A South African perspective. Small Rumin. Res. 150, 87-92.

Mdladla, K., Dzomba, E.F. & Muchadeyi, F.C., 2018. Landscape genomics and pathway analysis to understand genetic adaptation of South African Indigenous goat populations. Heredity. 120(4), 369-378.

Megersa, B., Markemann, A., Angassa, A., Ogutu, J.O., Piepho, H.P. & Zaráte, A.V., 2014. Impacts of climate change and variability on cattle production in southern Ethiopia: Perceptions and empirical evidence. Agr. Syst. 130, 23-34.

Meza, I., Rezaei, E.E., Siebert, S., Ghazaryan, G., Nouri, H., Dubovyk, O., Gerdener, H., Herbert, C., Kusche, J., Popat, E. & Rhyner, J., 2021. Drought risk for agricultural systems in South Africa: Drivers, spatial patterns, and implications for drought risk management. Sci. Tot. Environ. 799, 149505.

Mi, H. & Thomas, P., 2009. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. Methods Mol. Biol. 563, 123-40.

Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X. & Thomas, P.D., 2019. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). Nat. Protoc. 14, 703-721.

Michael, J.D., Baruselli, P.S. & Campanile, G., 2019. Influence of nutrition, body condition, and metabolic status on reproduction in female beef cattle: A review. Theriogenology. 125, 277-284.

Microsoft Corporation, 2018. *Microsoft Excel*, Available at: https://office.microsoft.com/excel

Mostert, B.E. & Exley, C.A.J., 2000. Inbreeding of the Bonsmara cattle breed. S. Afr. J. Anim. Sci. 30, 89-90.

Mucina, L., Rutherford, M.C. & Powrie, L.W., 2006. Biomes and bioregions of southern Africa. In: The vegetation of South Africa, Lesotho and Swaziland. In: L. Mucina, L. & M.C. Rutherford, SANBI, Pretoria. 31-51.

Nakanishi, T., Tokunaga, T., Ishida, T., Kobayashi, I., Katahama, Y., Yano, A., Erickson, L. & Kawahara, S., 2019. Changes in expression of the autophagy-related genes microtubule-associated protein 1 light chain 3β and autophagy related 7 in skeletal muscle of fattening Japanese Black cattle: a pilot study. Asian Austral. J. Anim. 32(4), 592.

Nardone, A., Ronchi, B., Lacetera, N., Ranieri, M. S. & Bernabucci, U., 2010. Effects of climate changes on animal production and sustainability of livestock systems. Livest. Sci. 130. 57–69.

NCBI. FKBP15, gene ID: 23307. Provided by GOA, 29 March, 2023. Accessed: 7 April 2023a. https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=23307.

NCBI. TMEM132B, gene ID: 114795. Provided by GOA, 29 March, 2023. Accessed: 7 April 2023b. https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=114795.

Neser, F.W., Erasmus, G.J. & Scholtz, M.M., 2008. The use of a cluster analysis in across herd genetic evaluation for beef cattle. S. Afr. J. Anim. Sci. 38(1), 51-57.

Nguyen, D.V., Nguyen, O.C. & Malau-Aduli, A.E., 2021. Main regulatory factors of marbling level in beef cattle. Veterinary and Animal Science. 14, 100219.

Nocek, J.E., Kautz, W.P., Leedle, J.A.Z. & Allman, J.G., 2002. Ruminal supplementation of direct-fed microbials on diurnal pH variation and in situ digestion in dairy cattle. J. Diary Sci. 85(2), 429-433.

Oishi, M., Gohma, H., Hashizume, K., Taniguchi, Y., Yasue, H., Takahashi, S., Yamada, T. & Sasaki, Y., 2006. Early embryonic death-associated changes in genome-wide gene expression profiles in the fetal placenta of the cow carrying somatic nuclear-derived cloned embryo. Mol. Reprod. Dev. 73(4), 404-409.

Ouellet, V., Negrao, J., Skibiel, A.L., Lantigua, V.A., Fabris, T.F., Marrero, M.G., Dado-Senn, B., Laporta, J. & Dahl, G.E., 2021. Endocrine signals altered by heat stress impact dairy cow mammary cellular processes at different stages of the dry period. Animals. 11(2), 563.

Padodara, R.J. & Jacob, N., 2013. Climate change: Effect on growth of animals. Basic Research Journal of Agricultural Science and Review. 2, 85-90.

Pantherdb. *ADRA1A*, Alpha-1A adrenergic receptor. Accessed March 2023a. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000031632%7CUniProtKB%3DA0A3Q1MAT2

Pantherdb. *ATG7*, Ubiquitin-like modifier-activating enzyme. Accessed March 2023b. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000035827%7CUniProtKB%3DA0A3Q1LXA9

Pantherdb. *C2H2orf88*, Small membrane A-kinase anchor protein. Accessed March 2023c. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000026994%7CUniProtKB%3DQ3SZY8

Pantherdb. *CDH18*, Cadherin-18. Accessed March 2023d. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000037844%7CUniProtKB%3DQ08DJ5

Pantherdb. *CSMD3*, CUB and Sushi multiple domains 3. Accessed March 2023e. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000038281%7CUniProtKB%3DA0A3Q1LWF2

Pantherdb. *DOCK1*, Dedicator of cytokinesis 1. Accessed March 2023f. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000031890%7CUniProtKB%3DF1MIX6

Pantherdb. *FKBP15*, Peptidylprolyl isomerase. Accessed March 2023g. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000005116%7CUniProtKB%3DF1MNV8

Pantherdb. *GLG1*, Golgi apparatus protein 1. Accessed March 2023h. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000002303%7CUniProtKB%3DA0A3Q1LL67

Pantherdb. *GMPS*, Glutamine amidotransferase. Accessed March 2023i. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000013013%7CUniProtKB%3DA0A3Q1LUN6

Pantherdb. *GTDC1*, Glycosyltransferase-like domain-containing protein 1. Accessed March 2023j. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000001132%7CUniProtKB%3DQ08DA7

Pantherdb. *HIBCH*, 3-hydroxyisobutyryl-CoA hydrolase, mitochondrial. Accessed March 2023k. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000007787%7CUniProtKB%3DA0A3Q1LUU2

Pantherdb. *IMMP2L*, Mitochondrial inner membrane protease subunit 2. Accessed March 2023l. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000004398%7CUniProtKB%3DQ2KI92

Pantherdb. *KCNJ16*, Potassium inwardly rectifying channel subfamily J member 16. Accessed March 2023m. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000024903%7CUniProtKB%3DQ0VD28

Pantherdb. KRAB domain-containing protein. Accessed March 2023n. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000053131%7CUniProtKB%3DA0A3Q1LXZ5

Pantherdb. *MAP4K3*, Mitogen-activated protein kinase kinase kinase kinase 3. Accessed March 2023o. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000016442%7CUniProtKB%3DE1BJH5

Pantherdb. *MYO3B*, Myosin IIIB. Accessed March 2023p. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000003626%7CUniProtKB%3DA0A3Q1M4N6

Pantherdb. *PRKG1*, cGMP-dependent protein kinase 1. Accessed March 2023q. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000018404%7CUniProtKB%3DP00516

Pantherdb. *RFX8*, RFX family member 8, lacking RFX DNA binding domain. Accessed March 2023r. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000033998%7CUniProtKB%3DE1BAR6

Pantherdb. *RTN4RL1*, Reticulon 4 receptor like 1. Accessed March 2023s. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000012302%7CUniProtKB%3DE1BEI7

Pantherdb. *SGMS1*. Sphingomyelin synthase 1. Accessed March 2023t. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000011748%7CUniProtKB%3DF1MF31

Pantherdb. *SMC4*, Structural maintenance of chromosomes protein. Accessed March 2023u. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000005862%7CUniProtKB%3DE1BMZ9

Pantherdb. *SPTB*, Spectrin beta chain. Accessed March 2023v. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000004732%7CUniProtKB%3DF1MKE9

Pantherdb. *TMEM132B*, Transmembrane protein 132B. Accessed March 2023w. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000006626%7CUniProtKB%3DE1BKR8

Pantherdb. *TSC22D1*, TSC22 domain family protein 1. Accessed March 2023x. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000047739%7CUniProtKB%3DQ3MHL6

Pantherdb. *TTC12*, Tetratricopeptide repeat domain 12. Accessed March 2023y. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000010008%7CUniProtKB%3DE1BEU1

Pantherdb. *WDR25*, WD repeat domain 25. Accessed March 2023z. http://www.pantherdb.org/genes/gene.do?acc=BOVIN%7CEnsembl%3DENSBTAG00000007233%7CUniProtKB%3DA0A3Q1NBM5

Pariset, L., Joost, S., Gargani, M. & Valentini, A., 2012. Landscape genomics in livestock. Chp. 10, 189-216. InTech.

Park, I.H. & Min, S.K., 2017. Role of convective precipitation in the relationship between subdaily extreme precipitation and temperature. J. Climate. 30(23), 9527-9537.

Patterson, N., Price, A.L. & Reich, D., 2006. Population structure and eigenanalysis. Plos Genet. 2(12), e190.

Pienaar, J., Louw, A. & Jordaan, D., 2019. A study on the potential product development for the commercialization and value add to beef products.

Pirinen, M., 2023. GWAS 2: P-values in GWAS. University of Helsinki. https://www.mv.helsinki.fi/home/mjxpirin/GWAS_course/material/GWAS2.html.

Polley, H.W., Briske, D.D., Morgan, J.A., Wolter, K., Bailey, D.W. & Brown, J.R., 2013. Climate change and North American rangelands: trends, projections, and implications. Rangeland Ecol. Manag. 66(5), 493-511.

Prayaga, K.C., Corbet, N.J., Johnston, D.J., Wolcott, M.L., Fordyce, G. & Burrow, H.M., 2009. Genetics of adaptive traits in heifers and their relationship to growth, pubertal and carcass traits in two tropical beef cattle genotypes. Anim. Prod. Sci. 49(6), 413-425.

Pritchard, J.K., Stephens, M., Rosenberg, N.A. & Donnelly, P., 2000. Association mapping in structured populations. Am. J. Hum. Genet. 67(1), 170-181.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J. & Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81(3), 559-575.

Purfield, D.C., Evans, R.D. & Berry, D.P., 2020. Breed-and trait-specific associations define the genetic architecture of calving performance traits in cattle. J. Anim. Sci. 98(5), p.skaa151.

Qwabe, S.O., Maiwashe, A. & Muchadeyi, F.C., 2013. Evaluation of the BovineSNP50 genotyping array in four South African cattle populations. S. Afr. J. Anim. Sci. 43(1), 64-67.

Rahman, M.B., Schellander, K., Luceño, N.L. & Van Soom, A., 2018. Heat stress responses in spermatozoa: Mechanisms and consequences for cattle fertility. Theriogenology. 113, 102-112.

Rashamol, V.P., Sejian, V., Bagath, M., Krishnan, G., Archana, P.R. & Bhatta, R., 2020. Physiological adaptability of livestock to heat stress: an updated review. J. Anim. Behav. Biometeorol. 6(3), 62-71.

Raza, S.H.A., Khan, S., Amjadi, M., Abdelnour, S.A., Ohran, H., Alanazi, K.M., Abd El-Hack, M.E., Taha, A.E., Khan, R., Gong, C. & Schreurs, N.M., 2020. Genome-wide association studies reveal novel loci associated with carcass and body measures in beef cattle. Arch. Biochem. Biophys. 694, 108543.

Reason, C.J.C., 2001. Subtropical Indian Ocean SST dipole events and southern African rainfall. Geophys. Res. Lett. 28(11), 2225-2227.

Reason, C.J.C. & Rouault, M., 2005. Links between the Antarctic Oscillation and winter rainfall over western South Africa. Geophys. Res. Lett. 32(7).

Reding, J.J., 2020. A Genome-wide association study of fertility and maternal traits in South African Bonsmara cattle (Doctoral dissertation, University of Pretoria).

Rellstab, C., Gugerli, F., Eckert, A.J., Hancock, A.M. & Holderegger, R., 2015. A practical guide to environmental association analysis in landscape genomics. Mol. Ecol. 24(17), 4348-4370.

Rezende, F.M., Rodriguez, E., Leal-Gutiérrez, J.D., Elzo, M.A., Johnson, D.D., Carr, C. & Mateescu, R.G., 2021. Genomic approaches reveal pleiotropic effects in crossbred beef cattle. Front. Genet. 12, 627055.

Rojas-Downing, M. M., Nejadhashemi, A. P., Harrigan, T. & Woznicki, S. A., 2017. Climate change and livestock: Impacts, adaptation, and mitigation. Climate Risk Management. 16, 145–163.

RStudio 2022.12.0 Build 353, ©2009-2022. Posit Software, PBC.

Rust, J.M. & Rust, T., 2013. Climate change and livestock production: A review with emphasis on Africa. S. Afr. J. Anim. Sci. 43(3), 255-267.

Saatchi, M., McClure, M.C., McKay, S.D., Rolf, M.M., Kim, J., Decker, J.E., Taxis, T.M., Chapple, R.H., Ramey, H.R., Northcutt, S.L., Bauck, S., Woodward, B., Dekkers, J.C., Fernando, R.L., Schnabel, R.D., Garrick, D.J. & Taylor, J.F., 2011. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. Genet. Sel. Evol. 43(40).

Sammad, A., Luo, H., Hu, L., Zhu, H. & Wang, Y., 2022. Transcriptome reveals granulosa cells coping through redox, inflammatory and metabolic mechanisms under acute heat stress. Cells. 11(9), 1443.

Santiago, G.G., Siqueira, F., Cardoso, F.F., Regitano, L.C.A., Ventura, R., Sollero, B.P., Souza, M.D., Mokry, F.B., Ferreira, A.B.R. & Torres, R.A.A., 2017. Genomewide association study for production and meat quality traits in Canchim beef cattle. J. Anim. Sci. 95, 3381-3390.

Santulli, G., Lombardi, A., Sorriento, D., Anastasio, A., Del Giudice, C., Formisano, P., Béguinot, F., Trimarco, B., Miele, C. & Iaccarino, G., 2012. Age-related impairment in insulin release: the essential role of β2-adrenergic receptor. Diabetes. 61(3), 692-701.

Saravanan, K.A., Panigrahi, M., Kumar, H., Bhushan, B., Dutt, T. & Mishra, B.P., 2020. Selection signatures in livestock genome: A review of concepts, approaches and applications. Livest. Sci. 241, 104257.

Scanzano, A. & Cosentino, M., 2015. Adrenergic regulation of innate immunity: a review. Front. Pharmacol. 6, 171.

Schober, P., Boer, C. & Schwarte, L.A., 2018. Correlation coefficients: appropriate use and interpretation. Anesth. Analg. 126(5), 1763-1768.

Serranito, B., Cavalazzi, M., Vidal, P., Taurisson-Mouret, D., Ciani, E., Bal, M., Rouvellac, E., Servin, B., Moreno-Romieux, C., Tosser-Klopp, G. & Hall, S.J., 2021. Local adaptations of Mediterranean sheep and goats through an integrative approach. Sci. Rep-UK. 11(1), 21363.

Sharma, M., Kambadur, R., Matthews, K.G., Somers, W.G., Devlin, G.P., Conaglen, J.V., Fowke, P.J. & Bass, J.J., 1999. Myostatin, a transforming growth factor-β superfamily member, is

expressed in heart muscle and is upregulated in cardiomyocytes after infarct. J. Cell. Physiol. 180(1), 1-9.

Sheng, H., Guo, Y., Zhang, L., Zhang, J., Miao, M., Tan, H., Hu, D., Li, X., Ding, X., Li, G. & Guo, H., 2021. Proteomic Studies on the Mechanism of Myostatin Regulating Cattle Skeletal Muscle Development. Front. Genet. 12, 752129.

Sherman, E.L., Nkrumah, J.D. & Moore, S.S., 2010. Whole genome single nucleotide polymorphism associations with feed intake and feed efficiency in beef cattle. J. Anim. Sci. 88(1), 16-22.

Schoonmaker, J. & Eastridge, M., 2013. Effect of maternal nutrition on calf health and growth. In: The proceedings of tri-state dairy nutrition conference. 63-80.

Scholtz, M.M., McManus, C., Leeuw, K.J., Louvandini, H., Seixas, L., Melo, C.D., Theunissen, A. & Neser, F.W.C., 2013. The effect of global warming on beef production in developing countries of the southern hemisphere. Nat. Sci. 5(1A), 14.

Shryock, D.F., Washburn, L.K., DeFalco, L.A. & Esque, T.C., 2021. Harnessing landscape genomics to identify future climate resilient genotypes in a desert annual. Mol. Ecol. 30(3), 698-717.

Silva, T. L., Gondro, C., Fonseca, P. A. S., da Silva, D. A., Vargas, G., Neves, H. H. R., Carvalho Filho, I., Teixeira, C. S., Albuquerque, L. G. & Carvalheiro, R., 2022. "Genetic mechanisms underlying feet and legs malformation in Nellore cattle: prioritization of GWAS results." In: Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP) Technical and species orientated innovations in animal breeding, and contribution of genetics to solving societal challenges. Wageningen Academic Publishers. 1598-1601.

Singh, M.S., Warren, R.A. & Jakob, C., 2019. A steady-state model for the relationship between humidity, instability, and precipitation in the tropics. J. Adv. Model Earth Sy. 11(12), 3973-3994.

Skibiel, A.L., Peñagaricano, F., Amorín, R., Ahmed, B.M., Dahl, G.E. & Laporta, J., 2018. In utero heat stress alters the offspring epigenome. Sci. Rep-UK. 8(1), 14609.

Smith, J. M. & Haigh, J., 1974. The hitch-hiking effect of a favourable gene. Genet. Res. 23, 23–35.

Solis, J.C., Byers, F.M., Schelling, G.T., Long, C.R. & Greene, L.W. 1988. Maintenance requirements and energetic efficiency of cows of different breed types. J. Anim. Sci. 66, 764-773.

South African Weather Service, 2020. Annual State of the Climate of South Africa 2019. Pretoria. South Africa.

South African Weather Service, 2021. Annual State of the Climate of South Africa 2020. Pretoria. South Africa.

South African Weather Service, 2022. Annual State of the Climate of South Africa 2021. Pretoria. South Africa.

Steyn, Y., van Marle-Köster, E. & Theron, H.E., 2014. Residual feed intake as selection tool in South African Bonsmara cattle. Livest. Sci. 164, 35-38.

Storfer, A., Patton, A. & Fraik, A. K., 2018. Navigating the Interface Between Landscape Genetics and Landscape Genomics. Front. Genet. 9(68).

Strydom, P.E., 2008. Do indigenous Southern African cattle breeds have the right genetics for commercial production of quality meat?. Meat Sci. 80(1), 86-93.

Strydom, P., Lühl, J., Kahl, C. & Hoffman, L.C., 2016. Comparison of shear force tenderness, drip and cooking loss, and ultimate muscle pH of the loin muscle among grass-fed steers of four major beef crosses slaughtered in Namibia. S. Afr. J. Anim. Sci. 46(4), 348-359.

Stucki, S., Orozco-terWengel, P., Forester, B.R., Duruz, S., Colli, L., Masembe, C., Negrini, R., Landguth, E., Jones, M.R., Nextgen Consortium, & Bruford, M.W., 2017. High performance computation of landscape genomic models including local indicators of spatial association. Mol Ecol. Resour. 17(5), 1072-1089.

Summer, A., Lora, I., Formaggioni, P. & Gottardo, F., 2019. Impact of heat stress on milk and meat production. Anim. Front. 9(1), 39-46.

Tainton, N.M. 1999. Production characteristics of the main grazing lands of South Africa. In: Veld management in South Africa. Ed. Tainton N.M. Pietermaritzburg: University of Natal Press, South Africa.

Taye, M., Kim, J., Yoon, S.H., Lee, W., Hanotte, O., Dessie, T., Kemp, S., Mwai, O.A., Caetano-Anolles, K., Cho, S. & Oh, S.J., 2017. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. BMC Genet. 18(1), 1-14.

Taylor, G., Swanepoel, F.J.C., Webb, E.C. & Stroebel, A. 2008. Effect of heifer frame size on their subsequent reproductive performance and pre-weaning performance of their calves, Aust. J. Exp. Agric. 48, 945-949.

Thonney, M. L., 2015. Genetics of Growth and Body Composition. The Genetics of Cattle, 2nd Edn. 523-543.

Thornton, P.K., Herrero, M.T., Freeman, H.A., Okeyo Mwai, A., Rege, J.E.O., Jones, P.G. & McDermott, J.J., 2007. Vulnerability, climate change and livestock-opportunities and challenges for the poor. Journal of Semi-Arid Tropical Agricultural Research.

Tolemariam, T., Adugna, E. & Fikadu, M., 2015. Development agents' perception about the effect of climate risk on livestock production. Livestock Research for Rural Development. 27(98).

Tucker, H. A. & W. D. Oxender, 1980. Seasonal aspects of reproduction, growth and hormones in cattle and horses. Prog. Reprod. Biol. 5, 155-180.

UP George Bonsmaras. About. Accessed April 2023. https://www.upgeorgebonsmaras.co.za/about

Urrutia, R., 2003. KRAB-containing zinc-finger repressor proteins. Genome Biol. 4(10), 1-8.

van der Walt, A.J. & Fitchett, J.M., 2021. Exploring extreme warm temperature trends in South Africa: 1960–2016. Theor. Appl. Climatol. 143(3-4), 1341-1360.

Van der Westhuizen, B., 2018. Genomic selection service for Bonsmara a reality. Die Rooi Ras. 2(1), 45.

Van der Westhuizen, J., 2019. Functional efficiency revisited Using one of the main pillars when selecting productive cattle. Nguni Journal. 19, 94-96.

Van der Westhuizen, J., Van der Westhuizen, B., Theron, H., Jordaan, F. & Maiwashe, N. 2011. The fitting of new comprehensive models for breeding value predictions: The Bonsmara breed at the fore font of science. Bonsmara Joernaal.

Van Graan, A.C., Neser, F.W.C., Van der Westhuizen, J. & Theron, H.E., 2004. Genetic and phenotypic parameter estimates of gestation length and birth weight in Bonsmara cattle. S. Afr. J. Anim. Sci. 34(2), 95-97.

van Marle-Köster, E. & Visser, C., 2018. Genetic improvement in South African livestock: can genomics bridge the gap between the developed and developing sectors?. Front. Genet. 9, 331.

van Marle-Köster, E., Visser, C., Sealy, J. & Frantz, L., 2021. Capitalizing on the potential of South African indigenous beef cattle breeds: a review. Sustainability-Basel. 13(8), 4388.

Van Rooyen, N. 2002. Veld management in the savannas. In: Game ranch management Ed. Bothma, J. du. P. 4th edition, Pretoria: Van Schaik, South Africa.

Vajana, E., 2017. Exploring livestock evolutionary history, diversity, adaptation and conservation through landscape genomics and ecological modelling. Doctorate. Doctoral School on the Agro-Food System.

Vajana, E., Barbato, M., Colli, L., Milanesi, M., Rochat, E., Fabrizi, E., Mukasa, C., Del Corvo, M., Masembe, C., Muwanika, V. B. & Kabi, F., 2018. Combining landscape genomics and ecological modelling to investigate local adaptation of indigenous Ugandan cattle to East Coast fever. Front. Genet. 9(385).

Valente, T.S., Baldi, F., Sant'Anna, A.C., Albuquerque, L.G., and Paranhos da Costa, M.J.R., 2016. Genome-wide association study between single nucleotide polymorphisms and flight speed in Nellore cattle. Plos One. 11(6), e0156956.

Vallejo-Trujillo, A., Kebede, A., Lozano-Jaramillo, M., Dessie, T., Smith, J., Hanotte, O. and Gheyas, A.A., 2022. Ecological niche modelling for delineating livestock ecotypes and exploring environmental genomic adaptation: The example of Ethiopian village chicken. Front. Ecol. Evol. 10, 866587.

Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N.H., Prestegard, J.H. & Schnaar, R.L., 2022. Essentials of Glycobiology [internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Vignal, A., Milan, D., Sancristobal, M. & Eggen, A., 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genet. Sel. Evol. 34, 275- 305.

Visagie, P.C., 2012. Effect of the production environment on the production efficiency of Bonsmara cows in South Africa. MSc Dissertation. 1-103.

Visser, C., van Marle-Köster, E., Myburgh, H.C. & De Freitas, A., 2020. Phenomics for sustainable production in the South African dairy and beef cattle industry, Anim. Front. 10(2), 12–18.

Visser, C., Lashmar, S.F., Reding, J., Berry, D.P. & van Marle-Köster, E., 2023. Pedigree and genome-based patterns of homozygosity in the South African Ayrshire, Holstein, and Jersey breeds. Front. Genet. 14.

Walker, C.E. & Drouillard, J.S., 2012. Effects of catecholamines on gut microflora and potential for beta-adrenergic agonists to impact ruminal fermentation. The Open Agricultural Journal. 6(1).

Webb, E. C., Visagie, P. C., van der Westhuizen, J. & Snyman, H. A., 2017. Influence of bioregion and environmental factors on the growth, size, and reproduction of Bonsmara cows. S. Afr. J. Anim. Sci. 47(4). 542-552.

Webber, H., Ewert, F., Olesen, J.E., Müller, C., Fronzek, S., Ruane, A.C., Bourgault, M., Martre, P., Ababaei, B., Bindi, M. & Ferrise, R., 2018. Diverging importance of drought stress for maize and winter wheat in Europe. Nat. Commun. 9(1), 4249.

Weber, K.L., Thallman, R.M., Keele, J.W., Snelling, W.M., Bennett, G.L., Smith, T.P., McDaneld, T.G., Allan, M.F., Van Eenennaam, A.L. & Kuehn, L.A., 2012. Accuracy of genomic breeding values in multibreed beef cattle populations derived from deregressed breeding values and phenotypes. J. Anim. Sci. 90, 4177–4190.

Weigand, H. & Leese, F., 2018. Detecting signatures of positive selection in non-model species using genomic data. Zool. J. Linn. Soc-Lond. 184(2), 528-583.

Weldon, D. & Reason, C.J.C., 2014. Variability of rainfall characteristics over the South Coast region of South Africa. Theor. Appl. Climatol. 115, 177-185.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H., 2019. Welcome to the tidyverse. J. Open Source Software. 4(43), 1686. doi:10.21105/joss.01686.

Wiener, P., Robert, C., Ahbara, A., Salavati, M., Abebe, A., Kebede, A., Wragg, D., Friedrich, J., Vasoya, D., Hume, D.A. & Djikeng, A., 2021. Whole-genome sequence data suggest environmental adaptation of Ethiopian sheep populations. Genome Biol. Evol. 13(3), evab014

Wilke, C.O., & Wiernik, B.M., 2022. Package 'ggtext'. https://wilkelab.org/ggtext/.

Wolfenson, D. & Roth, Z., 2019. Impact of heat stress on cow reproduction and fertility. Anim. Front. 9(1), 32-38.

Wright, J., Landwehr, G. & Chartan, E., 2019. Assessing the value of improved variable renewable energy forecasting accuracy in the South African power system. J. Energy South Afr. 30(2), 98-109.

Xu, Y., Cao, J., Huang, S., Feng, D., Zhang, W., Zhu, X. & Yan, X., 2015. Characterization of tetratricopeptide repeat-containing proteins critical for cilia formation and function. Plos One. 10(4), e0124378.

Yadav, A. K., Tomar, S. S., Jha, A. K. & Singh, J., 2017. Importance of molecular markers in livestock improvement: a review. Int. J. Agric. Innov. Res. 5(4).

Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M., 2011. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88(1), 76-82.

Yoder, J.B., Stanton-Geddes, J., Zhou, P., Briskine, R., Young, N.D. & Tiffin, P., 2014. Genomic signature of adaptation to climate in Medicago truncatula. Genetics. 196(4), 1263-1275.

Zhang, H., Wang, Z., Wang, S. & Li, H., 2012. Progress of genome wide association study in domestic animals. J Anim Sci Biotechnol. 3(26).

Zhang, L., Liu, J., Zhao, F., Ren, H., Xu, L., Lu, J., Zhang, S., Zhang, X., Wei, C., Lu, G. & Zheng, Y., 2013. Genome-wide association studies for growth and meat production traits in sheep. Plos One. 8(6), e66569.

Zhao, F., McParland, S., Kearney, F., Du, L. & Berry, D.P., 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. Genet. Sel. Evol. 47(1), 1-12.

Zhuang, Z., Xu, L., Yang, J., Gao, H., Zhang, L., Gao, X., Li, J. & Zhu, B., 2020. Weighted single-step genome-wide association study for growth traits in Chinese Simmental beef cattle. Genes. 11(2), 189.

Zimmermann, M.J., Kuehn, L.A., Spangler, M.L., Thallman, R.M., Snelling, W.M. & Lewis, R.M., 2021. Breed and heterotic effects for mature weight in beef cattle. J. Anim. Sci. 99(8), 1-7.

Addenda

Addendum A:

Table A. IBD values for EC clusters in Figure 3.1 and Figure 3.2

BONZAFF0000XXXX3944	0.058038
BONZAFF0000XXXX4272	0.056866
BONZAFF0000XXXX3782	0.055947
BONZAFF0000XXXX4264	0.055802
BONZAFF0000XXXX0971	0.054395
BONZAFF0000XXXX1003	0.052723
BONZAFF0000XXXX0930	0.052513
BONZAFF0000XXXX7856	0.051612
BONZAFF0000XXXX4041	0.051221
BONZAFF0000XXXX7898	0.050138
BONZAFF0000XXXX4298	0.048713
BONZAFF0000XXXX7872	0.048585
BONZAFF0000XXXX3758	0.048217
BONZAFF0000XXXX4322	0.047892
BONZAFF0000XXXX4314	0.047836
BONZAFF0000XXXX3733	0.047587
BONZAFF0000XXXX0955	0.047486
BONZAFF0000XXXX9906	0.047278
BONZAFF0000XXXX3951	0.046806
BONZAFF0000XXXX9617	0.046477
BONZAFF0000XXXX5021	0.045301
BONZAFF0000XXXX3915	0.045156
BONZAFF0000XXXX3790	0.045047
BONZAFF0000XXXX3527	0.043064
BONZAFF0000XXXX5252	0.04213
BONZAFF0000XXXX3808	0.041291
BONZAFF0000XXXX6249	0.040585
BONZAFF0000XXXX3923	0.040101
BONZAFF0000XXXX3561	0.039533
BONZAFF0000XXXX1029	0.039024
BONZAFF0000XXXX1011	0.034373
BONZAFF0000XXXX3538	0.033226
BONZAFF0000XXXX3931	0.033192
BONZAFF0000XXXX3645	0.032417
BONZAFF0000XXXX4544	0.031889
BONZAFF0000XXXX4569	0.031738
BONZAFF0000XXXX1078	0.03058
BONZAFF0000XXXX3504	0.029684
BONZAFF0000XXXX3520	0.028817
BONZAFF0000XXXX3553	0.027712
BONZAFF0000XXXX4577	0.027206
BONZAFF0000XXXX4585	0.026927
BONZAFF0000XXXX0540	0.024375

BONZAFF0000XXXX8190	0.022356
BONZAFF0000XXXX4593	0.020035
BONZAFF0000XXXX3689	0.018411
BONZAFF0000XXXX4407	0.017415
BONZAFF0000XXXX4957	0.017139
BONZAFF0000XXXX3572	0.016602
BONZAFF0000XXXX4399	0.014329
BONZAFF0000XXXX3697	0.01373
BONZAFF0000XXXX4431	0.012656
BONZAFF0000XXXX3705	0.012167
BONZAFF0000XXXX3598	0.009019
BONZAFF0000XXXX3606	0.007599

Addendum B: LEA Script

```
library(LEA)
# Creation a the genotypic file: "genotypes.lfmm"
file("ECgeno.ped")
ped2lfmm("ECgeno.ped", "ecgeno.lfmm", force = TRUE)
ped2geno("ECgeno.ped", "ecgeno.geno", force = TRUE)
project = NULL
project = lfmm("ecgeno.lfmm",
        "EC_Htemp.env",
        K = 3,
        repetitions = 10,
        project = "new")
# compute adjusted p-values
p = lfmm.pvalues(project, K = 3)
pvalues = p$pvalues
alpha = 0.2
L = length(pvalues)
# Benjamini-Hochberg's algorithm:
w = which(sort(pvalues) < alpha * (1:L) / L)
candidates = order(pvalues)[w]
# GWAS significance test
par(mfrow = c(1,1))
hist(pvalues, col = "lightblue")
plot(-log10(pvalues), pch = 19, col = "blue", cex = .7)
options(max.print=3000)
View(candidates)
```

Addendum C: Manhattan plots and histogram p-value plot results from the GWAS built into LEA.

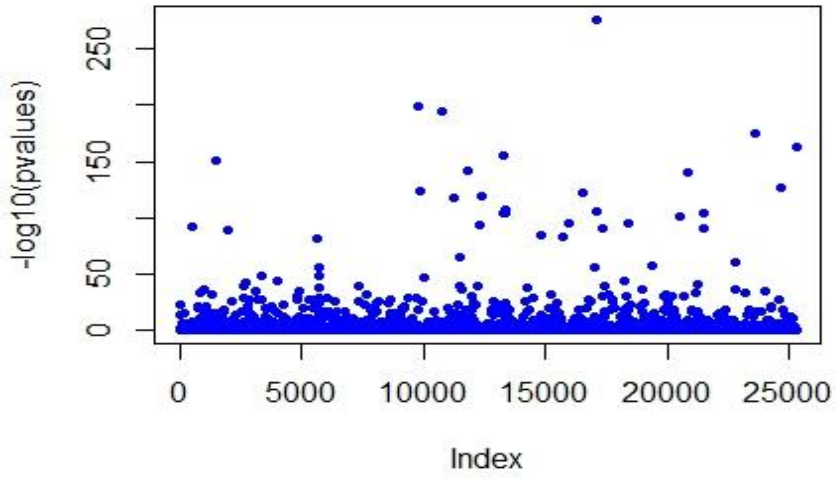


Figure C1. Manhattan plot for EC summer month temperatures, K = 3.

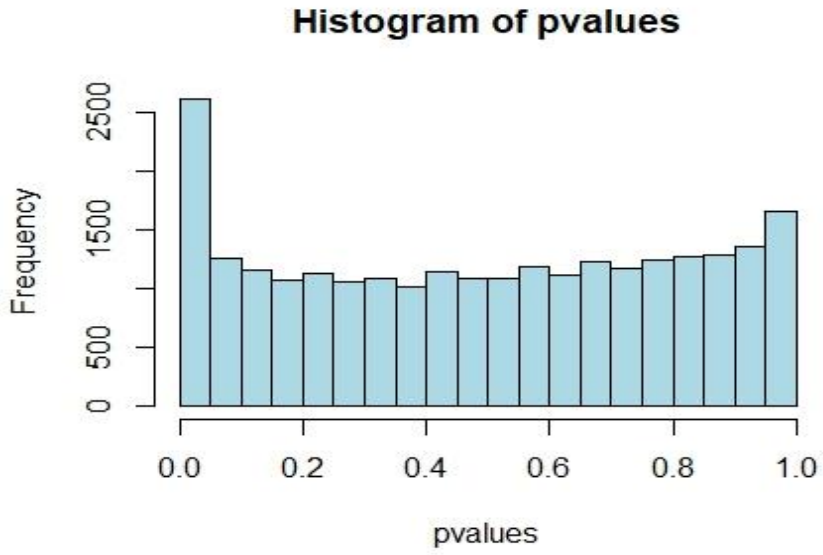


Figure C2. Pvalue histogram for EC summer month temperatures, K = 3.

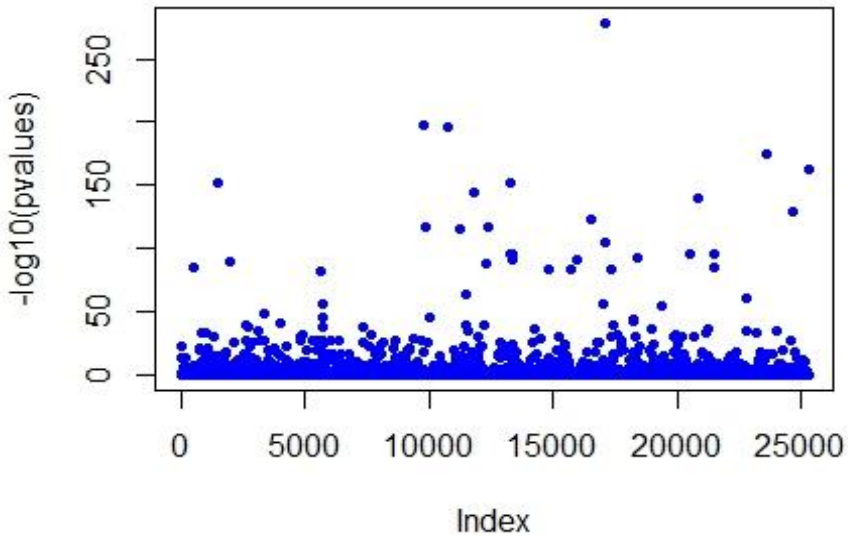


Figure C3. Manhattan plot for EC winter month temperatures, K = 3.

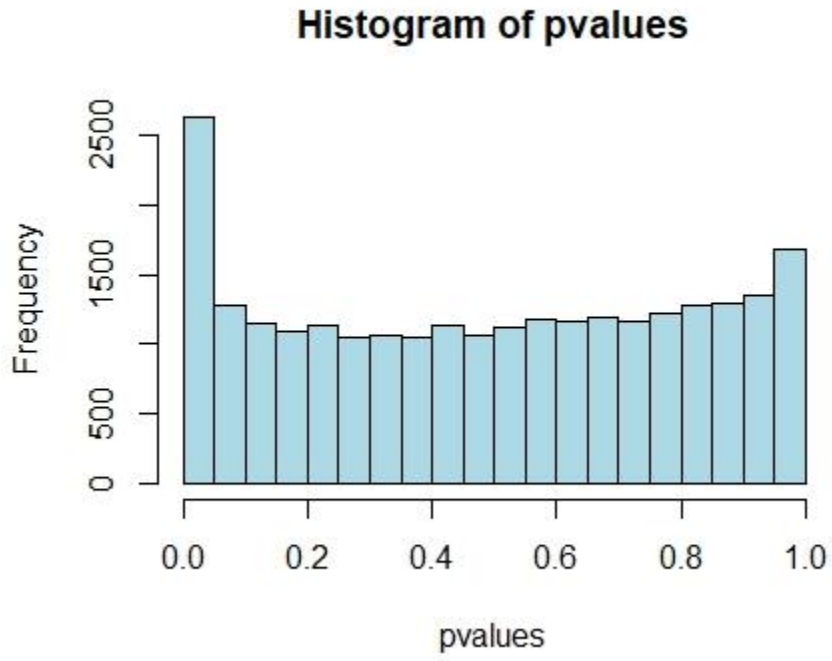


Figure C4. Pvalue histogram for EC winter month temperatures, K = 3.

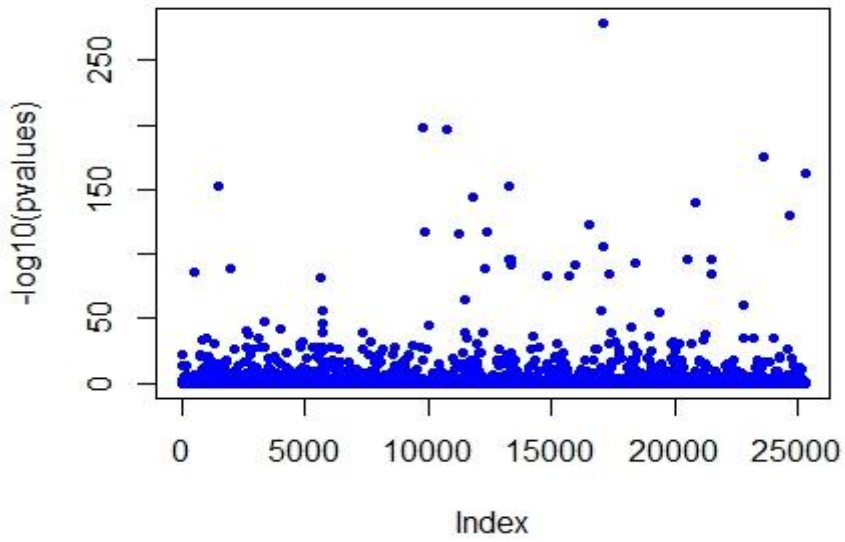


Figure C5. Manhattan plot for EC average annual precipitation, K = 3.

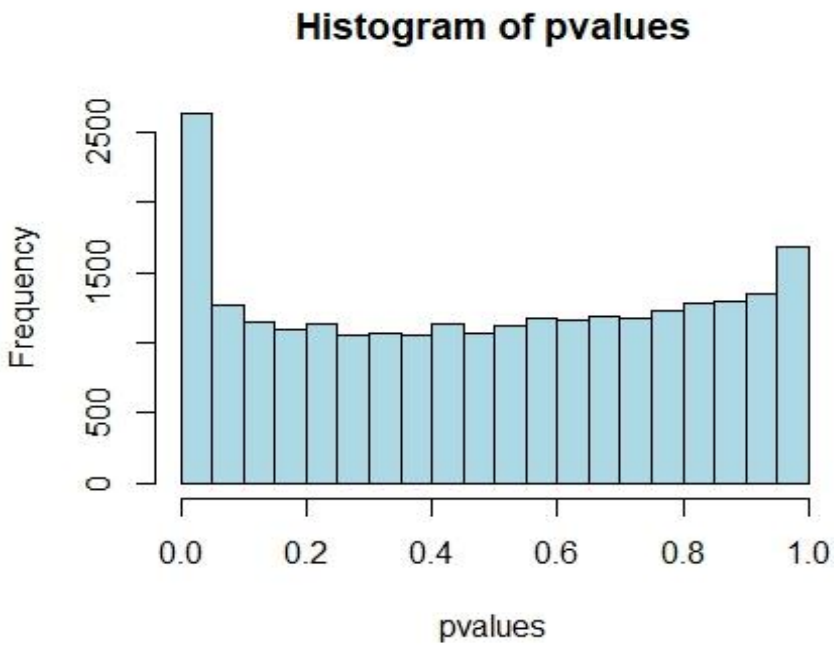


Figure C6. Pvalue histogram for EC average annual precipitation, K = 3.

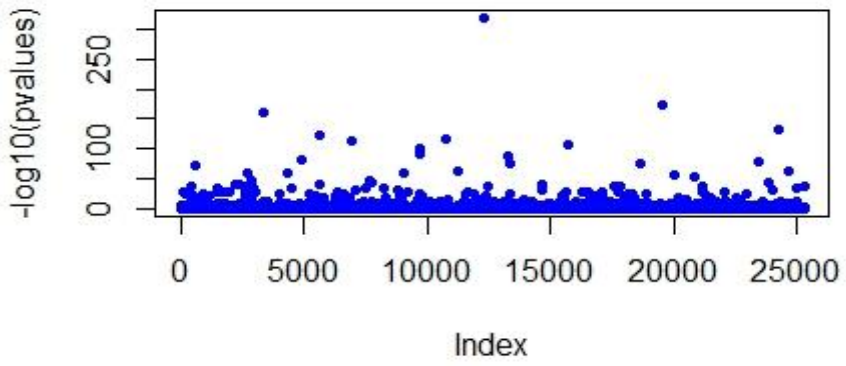


Figure C7. Manhattan plot for FS summer month temperatures, K = 1.

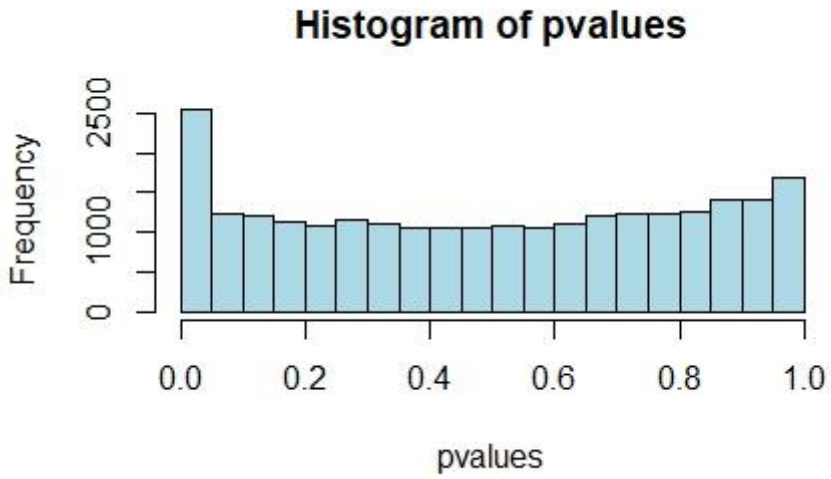


Figure C8. Pvalue histogram for FS summer month temperatures, K = 1.

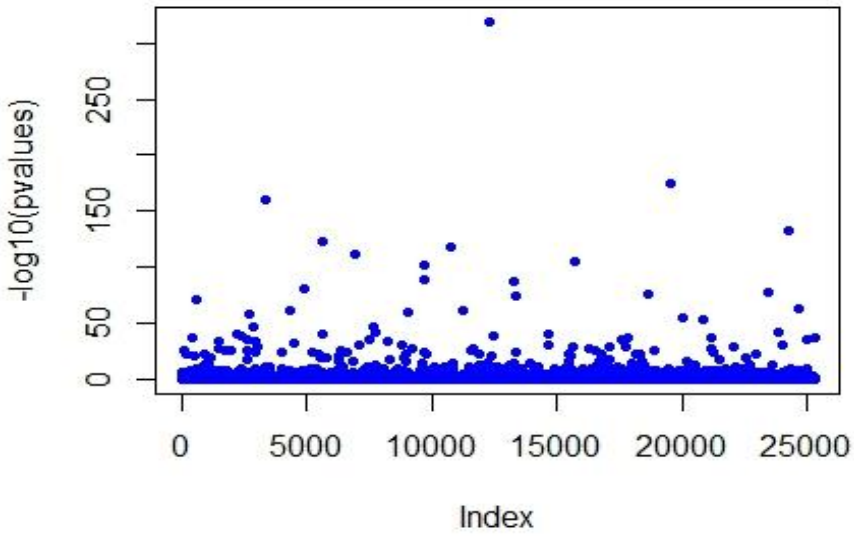


Figure C9. Manhattan plot for FS winter month temperatures, K = 1.

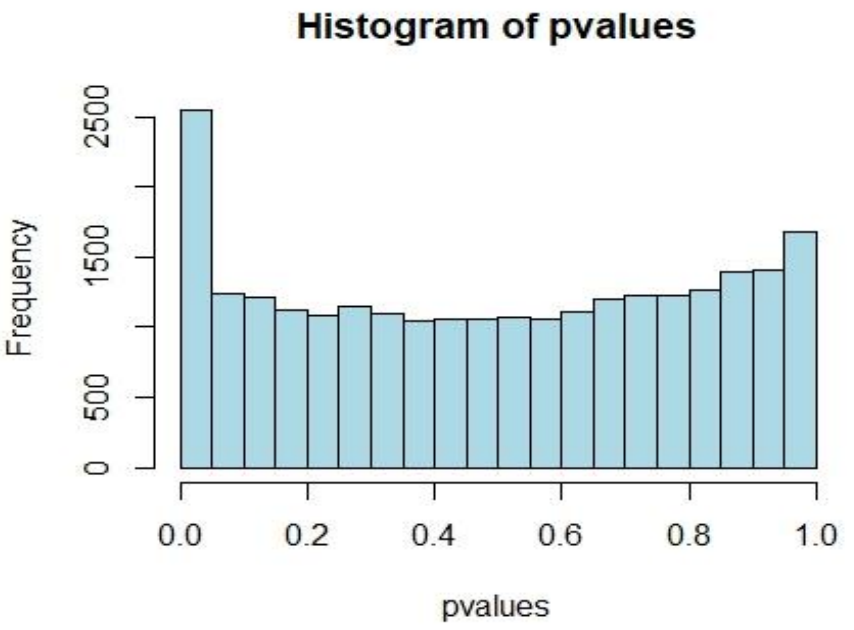


Figure C10. Pvalue histogram for FS winter month temperatures, K = 1.

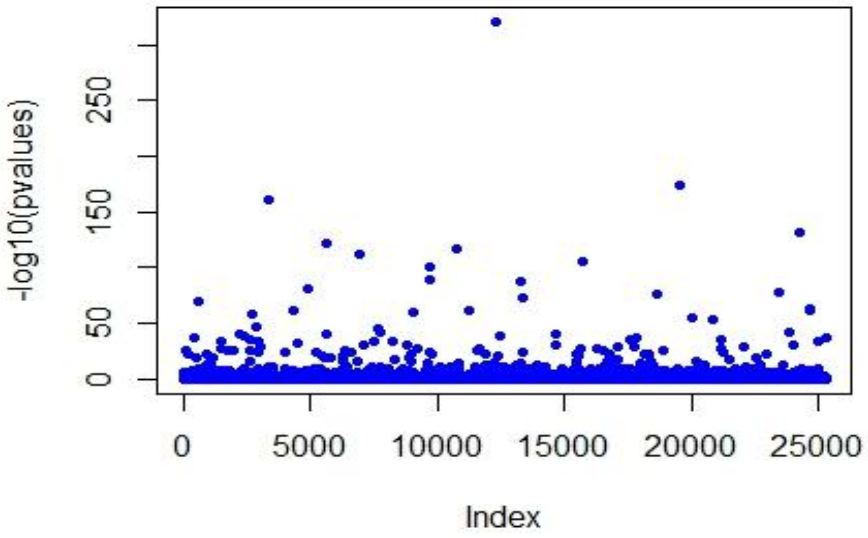


Figure C11. Manhattan plot for FS average annual precipitation, K = 1.

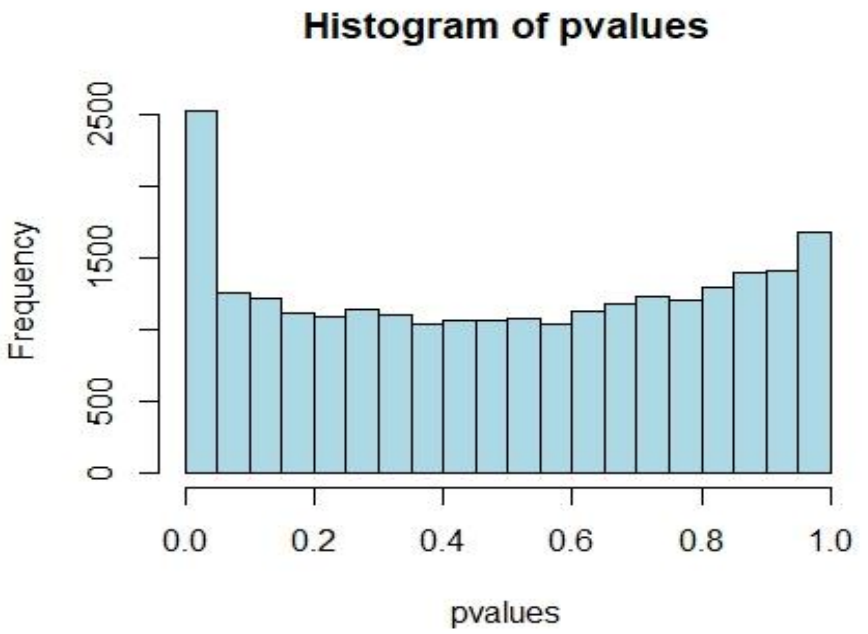


Figure C12. Pvalue histogram for FS average annual percipitation, K = 1.

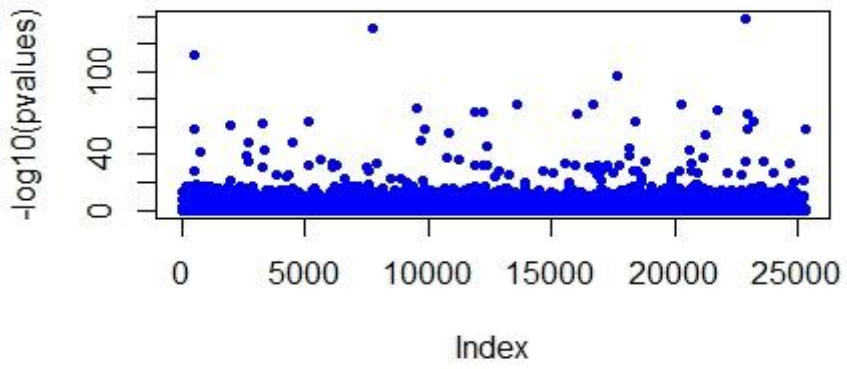


Figure C13. Manhattan plot for NW summer month temperatures, K = 1.

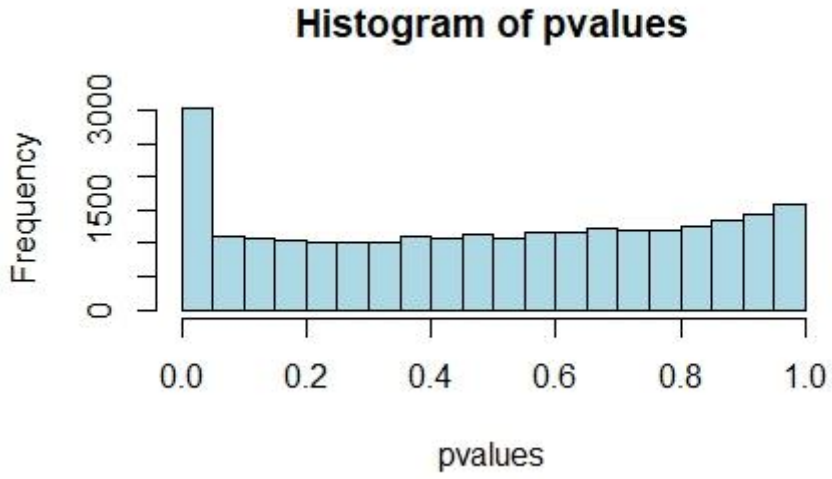


Figure C14. Pvalue histogram for NW summer month temperatures, K = 1.

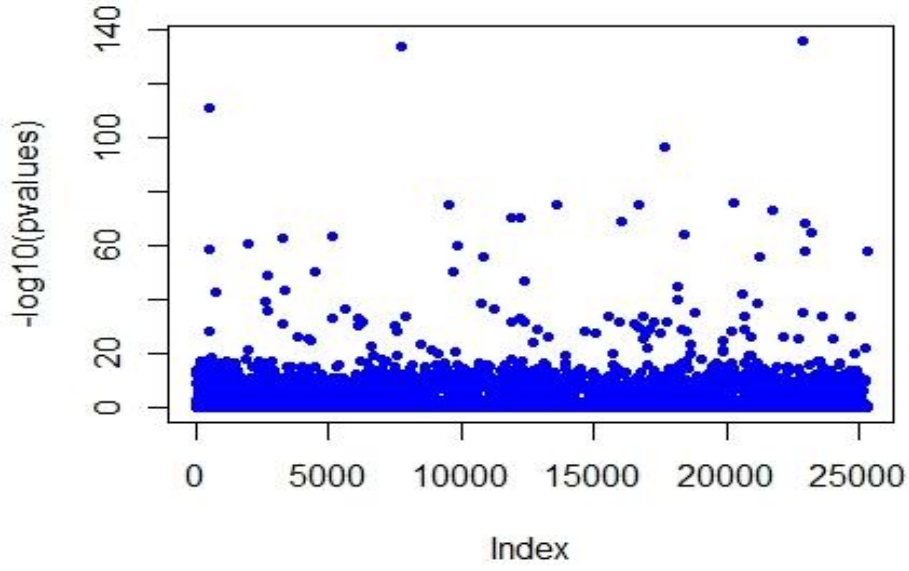


Figure C15. Manhattan plot for FS winter month temperatures, K = 1.

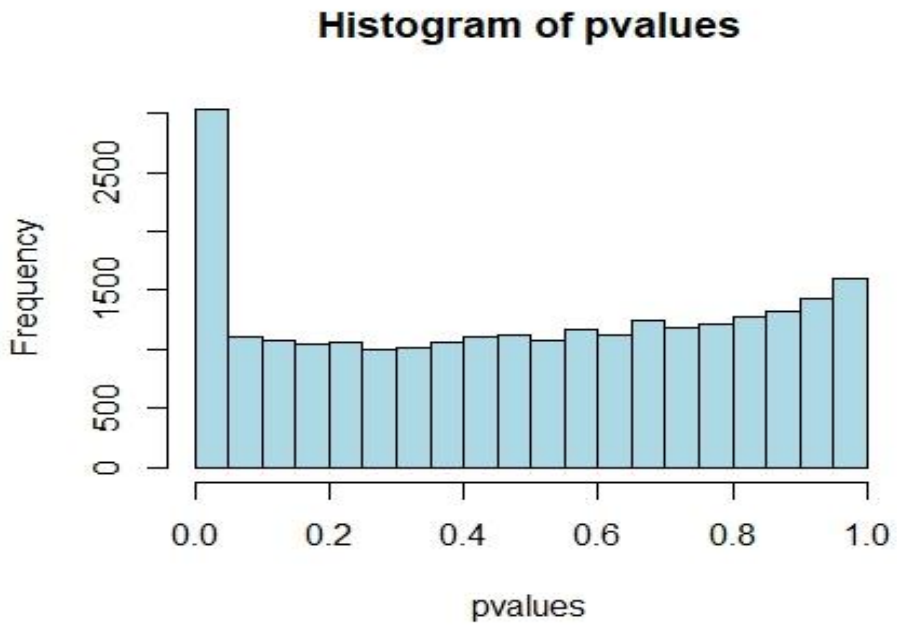


Figure C16. Pvalue histogram for NW winter month temperatures, K = 1.

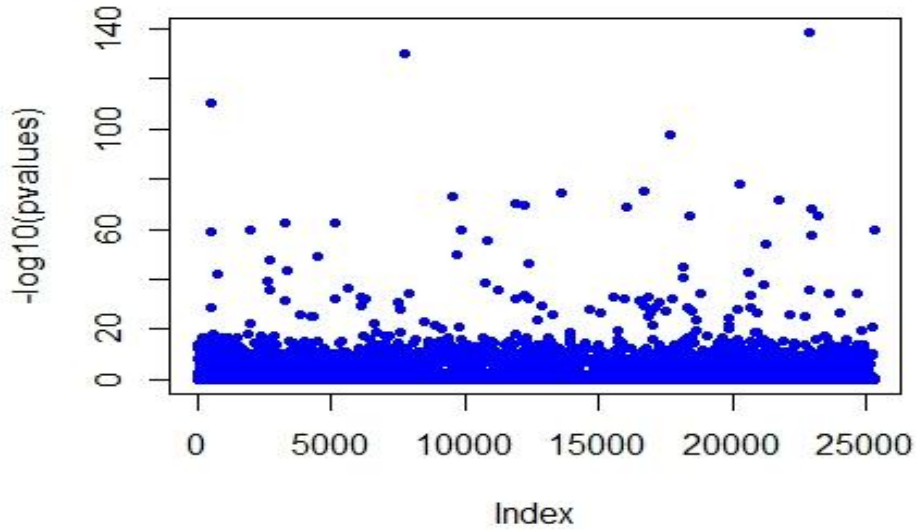


Figure C17. Manhattan plot for NW average annual precipitation, K = 1.

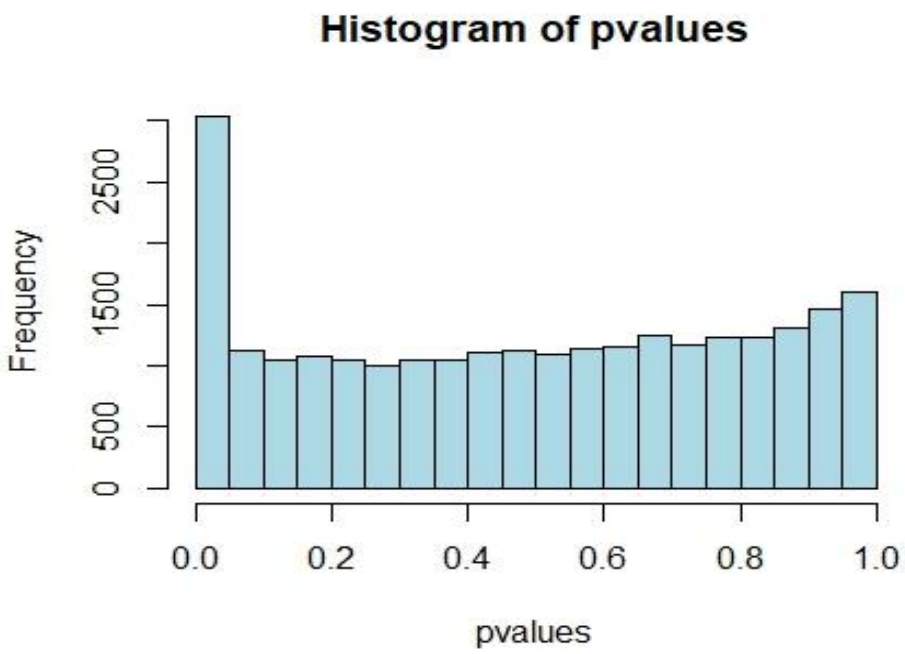


Figure C18. Pvalue histogram for NW average annual precipitation, K = 1.